STANFORD UNIVERSITY LIBRARIES
WITHDRAWN

MEASURING THE EFFECT OF FAMILY PLANNING PROGRAMS ON FERTILITY

Committee on Comparative Studies on Fertility and Fertility Planning

C. Chandrasekaran, Chairman
Gyorgy Acsádi
Mercedes B. Concepción
Ronald Freedman
Henri Leridon
Carmen Miró
Jean Morsa
Minoru Muramatsu

Sub-Committee on Demographic Impact of Family Planning Programs

C. Chandrasekaran, Convenor
William Brass
Henri Leridon
John Ross
Robert Potter
David Wolfers
Krishnamurthy Srinivasen

international union for the scientific study of population

The International Union for the Scientific Study of Population (IUSSP) is a non-profit, scientific organization composed of individual members. Established in 1928, it strives to stimulate research on population, develop interest in demographic matters, foster relations between people involved in population studies and disseminate scientific knowledge on population. It has consultative status with the Economic and Social Council of the United Nations and with Unesco.

Published for the Development Centre of the Organization for Economic Co-operation and Development, Paris.

MEASURING THE EFFECT OF FAMILY PLANNING PROGRAMS ON FERTILITY

Edited by

C. Chandrasekaran

and

Albert I. Hermalin

STANFORD UNIVERSITY LIBRARIES
WITHDRAWN

ORDINA EDITIONS
route de Goé 9
4830 Dolhain (Belgium)

HB891
M43
c.2

MEASURING THE EFFECT OF FAMILY PLANNING PROGRAMS ON FERTILITY

C. Chandrasekaran

and

Albert I. Hermalin

ORDINA EDITIONS
Route de Gos 2
4830 Dolhain (Belgium)

CONTENTS

FOREWORD, by Paul-Marc Henry XI

PREFACE, by C. Chandrasekaran, Albert I. Hermalin XIII

Chapter 1 INTRODUCTION, by W. Parker Mauldin, Gwendolyn Johnson-Acsádi 1

I. Development and Growth of Family Planning Programs 1
II. Views on the Efficacy of Family Planning Programs 4
III. Development of Evaluation Research 7
IV. Key Problems in Measuring the Effect of Family Planning Programs on Fertility 9
V. Scope of this Volume 11

Chapter 2 ACCEPTOR DATA, by C. Chandrasekaran 17

I. Purpose in Studying Acceptors 17
II. Concepts and Definitions Related to Acceptors 17
III. Sources and Procedures for Obtaining Data on Acceptors 21
IV. Use of Acceptor Data for Evaluating Program Performance 25
V. Evaluation of Program Progress Through Study of Trends of Acceptors 38
VI. Use of Acceptor Data for Improvement of Administration and Management 46
VII. Use of Acceptor Data for Assessing the Impact of Programs 50

Chapter 3 ACCEPTOR TARGETS, by John A. Ross 55

I. Introduction 55
II. Methods Used in Estimating Targets—Country Examples 59
III. Appraisal of Methods Used 82

Chapter 4 BIOSTATISTICS OF HUMAN REPRODUCTION, by Henri Leridon 93

I. Introduction: Microdemographic and Macrodemographic Approaches 93
II. Macrodemographic Background 95
III. The Physiological Framework of Reproduction 99
IV. Fecundability 103
V. Intra-Uterine Mortality 110
VI. The Physiological Basis of the Non-Susceptible Period 119
VII. Final Sterility 123
VIII. Introducing Contraception into the Model 127

Chapter 5 USE-EFFECTIVENESS OF CONTRACEPTION, by Robert G. Potter, Roger C. Avery 133

I. Introduction 133
II. Life-Table Approach 134
III. Methods of Extrapolation 143
IV. Extensions 150
V. Markov Chains 153

Chapter 6 BIRTHS AVERTED, by David Wolfers 163
I. Introduction 163
II. Concepts 165
III. Data: Methods of Handling, Sources and Limitations 176
IV. Illustrated Examples 195
V. Synthesis 208

Chapter 7 MATCHING STUDIES, by H. Bradley Wells 215
I. Introduction 215
II. Program Setting and Evaluation 215
III. Principles for Design of Controlled Trials 218
IV. Difficulties in Conducting Controlled Trials in Social Action Programs 219
V. Handling some Difficulties in Observational Studies of Social Action Programs 222
VI. Critique of Selected Matching Studies 236
VII. Summary and Discussion 241

Chapter 8 REGRESSION ANALYSIS OF AREAL DATA, by Albert I. Hermalin 245
I. Introduction 245
II. Structure and Rationale of Multivariate Areal Analysis 246
III. Examples of Areal Analysis in Program Evaluation 249
IV. Conceptual Framework and Specification 256
V. Sources and Nature of Data 260
VI. Choice of Areal Units and Spatial Autocorrelation 275
VII. Multivariate Techniques of Analysis 277
VIII. Interpretation and Use of Results 291

Chapter 9 RETROSPECTIVE SURVEYS, by K. Srinivasan 301
I. Introduction 301
II. An Overview of Purposes and Types of Surveys in Fertility and Related Fields 302
III. Objectives and Applications of Fertility/KAP Surveys 306
IV. Conceptual Schemes 311
V. Hypotheses 315
VI. Variables and Questionnaire 317
VII. Study Design 323
VIII. Analysis of Data 329
IX. Operational Problems and Non-Sampling Errors 333
X. Summary and Priority Areas for Further Research 338

Chapter 10 SIMULATION STUDIES, by Jane Menken 351
I. Introduction 351
II. Simulation Models 354
III. Applications of Simulation Models 363
IV. Discussion and Future Research 369
Appendix: Existing Models 373

Chapter 11 THE IMPACT ON FERTILITY OF SINGAPORE'S FAMILY PLANNING PROGRAM, by C. T. Chang, Philip M. Hauser 381
I. Introduction 381
II. The Family Planning Program 393
III. Response to the National Program 397
IV. Selected Features of the Program 404

V. The National Program in Relation to Fertility Changes 409
VI. Discussion and Conclusion 418

Chapter 12 THE IMPACT ON FERTILITY OF TAIWAN'S FAMILY PLANNING PROGRAM, by T. H. Sun 427
I. Introduction 427
II. Program Acceptances: Quantity and Quality 435
III. Use-Effectiveness 447
IV. Analysis of Fertility 468
V. Other Program Analysis 485
VI. Summary and Discussion 496

Chapter 13 OVERVIEW, by Albert I. Hermalin, C. Chandrasekaran 505
I. Introduction 505
II. Response and Use of Services 507
III. Duration and Effectiveness of Contraceptive Use 513
IV. Prevalence of Contraceptive Use 526
V. Births Averted Analysis 527
VI. Controlled Trials, Matching Studies, Areal Analysis and Surveys . . . 539
VII. Other Approaches 548
VIII. Conclusions 551

CONTRIBUTORS 555

INDEX 557

FOREWORD

It has been a privilege for the Development Centre of the Organisation for Economic Cooperation and Development (OECD) to make possible the undertaking of this international project on the Effect of Family Planning Programs on Fertility. A major part of the OECD's work in population over the past few years has been concerned with population assistance, which has grown in one decade from the order of $10 million to $200 million in 1973, and of which some two thirds have been made up by aid for family planning programs. There is therefore a keen interest both among the donors of aid as well as those countries who have been massively investing in those programs to get some idea of the effectiveness of these efforts. This book should assist in the extremely difficult task of assessing program effectiveness. It represents a major step forward in the field of evaluation of family planning programs.

The Development Centre is convinced that the findings and the results of the various studies assembled in this work can be of immense value to decision-makers, particularly to countries with family planning programs. It is accordingly following up the publication of this book by special measures to bring them to the attention of the people in both developing countries and aid agencies to whom they may be most relevant. Meanwhile, the Development Centre would like to congratulate the IUSSP and the editors and authors of this volume on the completion of a complex, difficult, but extremely productive and useful project.

Paul-Marc Henry
President
Development Centre
OECD

PREFACE

This publication grew out of a meeting of experts convened by the International Union for the Scientific Study of Population (IUSSP) in July 1971 to consider the problems that arise in assessing the demographic impact of family planning programs.* The participants felt that considerable progress had been made in the last two decades in developing techniques for dealing with the planning, execution and evaluation of family planning programs. However, it became evident that such developments were taking place sporadically, and in the face of growing program needs, a more systematic approach to the relevant methodology was called for. The expert group felt that as a first step, efforts should be made to describe "the state of the art" in studying the effect on fertility of family planning programs. A joint plan was therefore devised, whereby the Development Centre of the Organisation for Economic Cooperation and Development (OECD) would finance the IUSSP in the development of a project along these lines, and which has resulted in the present publication.

This volume consists of thirteen chapters. The first provides an introduction and the last gives an overview of the current state of knowledge and highlights some of the research issues involved. Each of the nine chapters 2 through 10 deals with a topic of special interest in assessing the impact on fertility of family planning programs. Chapters 11 and 12 are case studies dealing respectively with the family planning programs of Singapore and Taiwan—two areas rich in statistical information and therefore likely to exemplify best what can be done to assess the impact of these programs.

The authors of chapters other than the overview worked from a common set of guidelines in preparing the outlines for their chapters. These outlines were discussed and modified at a meeting held in August 1972 ** with a view to achieving a consensus on the content

* The following persons attended the meeting: Francis E. Bland; William Brass; C. Chandrasekaran; L. P. Chow; Ronald Freedman; G. Z. Johnson-Acsádi; Henri Leridon; W. Parker Mauldin; Robert G. Potter; T. K. Ruprecht; K. Srinivasan; G. J. Stolnitz; S. M. Wishik and David Wolfers.

** The following persons attended the meeting: C. Chandrasekaran; Henri Leridon; W. Parker Mauldin; Jane Menken; Robert G. Potter; John Ross; K. Srinivasan; C. M. Suchindran and David Wolfers.

of each chapter and minimizing overlap of closely related subject matter across chapters. The draft of each chapter was reviewed by two advisors—generally the authors of other chapters—and revised according to their suggestions.

In so far as applicable, each chapter attempts to focus on the basic concepts associated with a technique or special topic and their relevance to the study of fertility change; the data related to these concepts and their sources; examples of the applications of each technique; limitations in terms of data or methods; suggested improvements in data collection or methods of analysis; and issues requiring further attention. Within this framework, the relative weight to be given to each of these items, their ordering, and the style of presentation were left to the authors in the editors' belief that the diversity of subject matter precluded a highly standardized format. Similarly, no attempt was made to interfere with the personal views of the authors. The responsibility for the statements made, whether technical or otherwise, is that of their respective authors.

Although each chapter deals with a separate topic, the chapters have been ordered in a manner designed to facilitate an understanding of the field, by starting with simpler concepts and techniques and moving to more complex ones. In fact, the ordering of the chapters follows somewhat closely the way in which ideas and methods in the field have developed during the last two decades.

This publication is designed to give technical orientation to demographers, statisticians and persons from other disciplines interested in evaluating the impact of family planning programs. It should serve as an introduction for graduate students undergoing training in this field. Research workers in the field of family planning evaluation will also, it is hoped, find this publication stimulating and suggestive of several new lines of endeavor. It is expected that this volume will be used as a reference by those who want to pursue in detail the techniques covered in a specific chapter, and as an intermediate level text or reader by those who want to cover all the issues presented. Both types of users may find it helpful to start with Chapters 1 and 13, which attempt to place each technique in a larger perspective and, in some cases, provide supplementary technical information or analysis.

The publication would not have been possible without the willing cooperation of a large number of people. Many of the ideas came from the participants in the meeting of experts in July 1971 and the meeting to discuss chapter outlines held in August 1972. In addition to the authors, the following persons assisted in reviewing chapters or in the preparation of material for the overview: William Brass;

L. P. Chow; Ronald Freedman; Louis Henry; Dorothy Nortman; Anrudh Jain; T. Paul Schultz; K. S. Srikantan; C. M. Suchindran; Christopher Tietze and K. Venkatacharya. Professors Jeanne Ridley and David Goldberg assisted in the reading of the final text. The responsibility of editing for style was undertaken by Barbara Segal, with the assistance of Susan Etter, and the typing work involved in preparing the text for the printers was shared by Virginia Aquino and Cecile Cuffley.

On behalf of the IUSSP and its Committee on Comparative Studies on Fertility and Fertility Planning, under whose auspices this project was undertaken, we wish to express our gratitude to the Development Centre of the OECD, which has supported this project from its inception and has provided the necessary funds for carrying it through. We would also like to thank the various staff of the Development Centre, present and past, who have given this project their support and encouragement; the United Nations Funds for Population Activities(UNFPA) for supplementary assistance and office facilities and its Executive Director, Mr. Rafael M. Salas, for his continued interest; and the IUSSP's Executive Secretary, Mr. Bruno Remiche, for his efforts to get the text published expeditiously.

C. Chandrasekaran
Albert I. Hermalin

Chapter 1
INTRODUCTION

by
W. Parker Mauldin and Gwendolyn Johnson-Acsádi *

I. DEVELOPMENT AND GROWTH OF FAMILY PLANNING PROGRAMS

DEFINITIONS OF FAMILY PLANNING AND FAMILY PLANNING PROGRAMS

The term *family planning* has a variety of usages. In this volume it refers to the deliberate effort of couples or individuals to regulate fertility by delaying or spacing births and/or limiting their numbers. It implies deliberate action to avoid conception or a live birth. Family planning has been practiced in one form or another throughout man's history, but it was not publicly advocated until early in the 19th century, and then only by private groups. In recent years, such advocacy has been, *inter alia*, in terms of formal government policy, implemented by family planning programs.

By *family planning programs* we mean organized efforts to provide birth control information and services on a voluntary basis to a target population, with a view to lowering fertility and/or improving health. Such programs are typically governmental in funding and administration and are usually undertaken as part of a wider set of activities aimed at fostering economic and social welfare at both the individual and societal levels. From this broader perspective, the reduction of fertility is but one of a number of important steps leading to a higher level of living.

While this volume is concerned with the spacing and limiting of births within the context of family planning programs, of which a primary aim is the spread of effective birth regulation practices, programs often have corollary objectives. These may be the reduction of maternal and child deaths, an improvement in child health,

* Gwendolyn Johnson-Acsádi is a staff member of the United Nations Population Division. The views reflected in this chapter are her own and do not necessarily represent those of the United Nations.

and a reduction in non-medical and illegal induced abortions. The programs provide information, service and supplies that will enable people to regulate fertility both in terms of completed family size and the spacing of children and, insofar as possible, to choose among the most effective, safe and acceptable means of fertility control. In addition, a program may include activities aimed at persuading couples to adopt services, or at reducing desired fertility, and some may involve incentives as part of this process.

A family planning program as such does not seek to change fertility by changing the social institutions and practices that directly or indirectly influence reproductive behavior, although such changes may occur as a by-product of the program. An example of a change in social institutions or practices that probably would have an appreciable effect upon fertility would be an alteration in attitudes towards the employment of women and markedly greater opportunities for them to find gainful employment. But a program whose primary purpose is to increase the employment of women should not be considered part of a family planning program, even though such an effort may have been motivated by the goal of fertility reduction.

DEVELOPMENT OF FAMILY PLANNING PROGRAMS

Family planning programs are a relatively recent phenomenon. Although national policy with respect to fertility dates back some centuries, it was not until 1951 that India adopted the first national program aimed at lowering birth rates and, hence, the rate of national population growth (United Nations, 1972: 65). Family planning did not, however, gain any appreciable acceptance as a policy instrument until the early 1960's, as a result, among other things, of growing awareness of the dampening effect of rapid population growth on economic development, and the availability of new contraceptive methods.

The countries of Asia took the lead in the family planning movement and, while some countries of Latin America have introduced programs, only a few African countries have done so. Further, with the exception of those in Asia, most countries have implemented the programs mainly for health, rather than demographic reasons. During the 1960's, twenty-two countries adopted official policies to reduce their population growth rates and, as instruments of this policy, either established national family planning programs or lent support to private agencies involved in family planning. In the same period, sixteen additional governments gave official support to family planning activities, but for reasons other than the reduction of fertility

and the rate of national population growth (Berelson, 1974). During 1970-73, eight more countries established family planning programs with the aim of achieving lower fertility, and twelve others began official support of family planning activities, although not specifying the reduction of population growth rates among their goals.

Of the countries seeking to lower their rates of population growth, seventeen have set demographic targets (Nortman, 1973) and several others have established targets for numbers of acceptors of family planning methods by year, often by method. Several factors have led to the revision of demographic targets over time: the end of the plan period has occasioned a new goal; better data have permitted more precision in stating the current level of fertility; failure to reach the stated goal has called for a restatement (typically a lowering of the target and/or an extension of the time period).

STRUCTURE AND MANAGEMENT OF FAMILY PLANNING PROGRAMS

Where programs are government sponsored, administration is typically connected with an official maternal and child health service. In some cases, however, separate government agencies have been specially established and accorded responsibility either for delivering services or for coordinating the work of two or more public and/or private agencies; in still other cases, governments have entrusted implementation of the family planning programs to private family planning associations. In addition to health ministries, public agencies having responsibility for family planning work include ministries of social welfare, ministries of labor, community delvelopment agencies and the army (Nortman, 1973: 34-44; Keeny, 1974; Berelson, 1974).

The components of a family planning program vary greatly from program to program, and over time, within a given program. Typical programs include: an informational component; service points at hospitals, health centers, special clinics, mobile clinics, offices of private physicians, and occasionally delivery of supplies by paraprofessional and lay personnel; a variety of contraceptive methods, usually free or at subsidized prices; facilities for male and female sterilization; and sometimes provision for induced abortion. Monetary incentives to field workers and to doctors are common, though by no means universal. Such incentives to clients or family planning acceptors are relatively rare (Rogers, 1971), though they have been used, particularly in vasectomy programs. They are currently being tried in two field experiments: a deferred bond scheme to workers in tea estates in India (Ridker, 1971) and an educational bond scheme

in Taiwan (Wang and Chen, 1973); in each experiment the amount of the bond increases over time if the woman avoids another birth. In theory at least, activities in the private sector can also form part of the overall program, and some efforts to involve the private sector more effectively have been reported in a few programs, such as those of Ghana and Colombia.

The structure and management of a program have a marked effect on the quality of service as well as on the range of birth control methods available. For example, a program can offer sterilization and abortion only if its participating staff includes trained medical personnel. Program administration is also of paramount importance in the achievement of program goals (Berelson, 1974). Together with logistics, it is the subject of continuous internal and external evaluation. However, these aspects are not of primary concern in this volume, and no explicit attention is given to the methodology of their evaluation.

II. VIEWS ON THE EFFICACY OF FAMILY PLANNING PROGRAMS

The initiation of any large-scale social program reflects the values of a society. Debate over the worth of a program can stem from differences about the means of achieving a goal or from the priorities to be attached to a set of values. Thus, family planning programs are in competition for scarce resources with other objectives in the society and with alternative strategies for achieving economic development and social welfare.

Any social action program, in addition, is based on a set of "validity assumptions" (Suchman, 1967: 42). As distinct from values, validity assumptions relate more narrowly to program objectives and the rationale behind fashioning a program along specific lines. These assumptions stem from our understanding of the process of fertility decline and involve, among others, components such as the following: a significant proportion of couples are having more children than desired, with this surplus arising in part from restricted knowledge and availability of family planning services, making such services more freely available and increasing the population's motivation to utilize them will reduce the number of unwanted children, and hence overall fertility. Controversies over the worth of a social action program are likely to be greater where the scientific underpinnings are open to question. Thus, concern about the efficacy of family planning programs reflects in part the lack of closure in our explanations for the historical fertility decline in currently indus-

trialized countries and of the current determinants of fertility cross-sectionally within countries.

Essentially there have been two viewpoints about the importance of family planning programs for fertility change. One minimizes their importance, maintaining that changes in social structure and economy are needed. In this view, specific development measures such as improved universal education might well prove a more fruitful way of lowering fertility (Davis, 1963; Hauser, 1967). This was the dominant view at the World Population Conference in Bucharest (Mauldin *et al.*, 1974). On the other hand, however, many proponents of the programs maintain that the majority of high fertility groups are desirous of controlling family size and would do so if provided with the knowledge and means (Bogue, 1964). An intermediate position, supported by many, is that the lag between the early stages of modernization and a shift to smaller family size desires can be shortened significantly by the introduction of family planning programs, and that this in turn will aid the development process.

It is generally acknowledged that fertility is more likely to decline within the context of economic and social advancement. Accordingly, with the modernization of society, there are changes in the structure and function of the family that make large numbers of children a liability rather than the asset they are in many traditional cultures (see, for example, Freedman, 1966: 41). One view is that the principal change responsible for the development of controlled fertility has been a change in attitudes or motivation, and social mobility has been hypothesized as a primary factor influencing this motivation (Berent, 1952; Girard, 1951). In addition, birth control is said to have been only one of the several responses to high fertility observed among the people of northwestern Europe and Japan, who reduced birth rates by postponement of marriage, increased celibacy, abortion, contraception and overseas migration (Davis, 1963).

A particularly relevant question then is whether in the developing countries, fertility regulation can be fostered through family planning in the absence of some or all of the modernization changes that preceded or accompanied fertility declines in the currently developed countries.

Freedman (1966) hypothesized that:

> "... family planning is unlikely to be widely adopted in a country until there has been a significant mortality decline and until there has been enough social and economic development to lessen dependence on local and familial institutions and to make smaller families more rewarding than larger families."

He noted further that once a certain threshold level of development

is reached, family planning programs could be successful in reducing fertility where it remains high. The "threshold" hypothesis holds that where fertility is initially high, it will be influenced very little, if at all, by early improvements in social and economic conditions and by family planning service programs, but that once a certain level of development has been achieved, fertility will decline (United Nations, 1964). Little is known, however, about the specific content and levels of the hypothesized threshold. For example, the decline of mortality, presumably a key factor in fertility control, now occurs in the absence of full or even moderate modernization (see, for example, Davis, 1963).

An alternative view is that economic and social advancement could benefit substantially from a reduction in rates of population growth. Where there is a will and adequate assistance, a family planning program could reduce fertility in newly developing countries to reasonably low levels within two or three decades (Notestein, 1967). This view was supported by Berelson (1969), among others, who held that family planning would proceed more rapidly in the context of modernization, but that well-designed and administered programs would afford at least a start at family planning among some segments of the population. Recent studies show that fertility has fallen in some societies in the absence of modernization, and one author states that "There is no convincing basis for asserting that a program of indoctrination in the advantages to health and welfare from reduced fertility would inevitably be a failure in a rural poorly educated population." (Coale, 1973.)

On the other hand, some, particularly traditional Marxists, hold that what we identify as problems of population growth and structure are economic problems stemming from the character of the economy; family planning could not, therefore, expedite development. In some developing countries, particularly in Africa, population growth is seen as a false issue, an attempt by rich countries to serve their own self-interest, misallocating international assistance to population control rather than to full development.

In a summary assessment of family planning programs early in 1974, Berelson wrote:

> "In general, it might be said that persuasion and enlightenment seem to be minimally effective in the absence of major social changes; that family planning programs are useful but limited in impact; that incentives in either direction may be somewhat effective but could become unduly expensive with large impacts; that broad social change, though effective, is difficult, slow, and very costly; and that coercion is morally repugnant, politically unacceptable, and virtually impossible of administration.
>
> Accordingly, deliberate efforts to reduce fertility take time to have

effect—and particularly since almost all these measures, all except the provision of family planning services, imply a difference between what the individual couple consider desirable by way of fertility behaviour and what the society (i.e. government) considers desirable. Substituting the latter for the former is naturally no easy task."

(Berelson, 1974)

III. DEVELOPMENT OF EVALUATION RESEARCH

Family planning programs were from the outset experimental, in that there were no guidelines for their organization and management and no assurance that they could accomplish the goals specified. Added to this uncertainty was the obligation of administrators to justify the resources invested in their programs, especially in the earlier years when family planning had not gained wide acceptance and financial resources were less assured. Further, organization and explicit goals varied from one program to another, so that each program was required to develop some practical and realistic methods of judging levels of performance and of ascertaining the rate of progress toward accomplishing stated aims.

Thus, with the development of family planning programs there have been periodic assessments of the nature of the programs and of the means of evaluating them (Berelson, 1970; Bean and Seltzer, 1968; Chandrasekaran and Freymann, 1965; Lapham and Mauldin, 1972; Mauldin, 1967 ; Reynolds, 1972 ; Seltzer, 1970; United Nations, 1972). At the same time there has been renewed attention to the general characteristics of evaluative research in connection with public service programs (Suchman, 1967 ; Caro, 1971; Rossi and Williams, 1972).

Evaluation has been defined in many ways, but most definitions agree that it involves the determination of the results achieved by a program directed toward some valued objective (Mauldin, 1968; Suchman, 1967; 31-32). Account should be taken of indirect as well as direct consequences of a program; of unexpected and undesirable effects as well as expected and desirable ones. The process of evaluation is an important component of program development and is particularly valuable in advancing new programs. In family planning programs, the information provided by evaluation helps administrators to set realistic and internally consistent goals with respect to program objectives; and through the assessment of achievement, helps determine the changes necessary in organization and/or performance to attain specified targets.

In a family planning program a series of objectives may be distinguished. Many of these can be ordered in hierarchical fashion

which progress from immediate objectives to the ultimate goal (Chandrasekaran and Freymann, 1965). Such a chain would involve objectives relating to the training and recruitment of personnel, the establishment of clinics, the production of contraceptives, the dissemination of information, the recruiting of acceptors and ultimately the reduction of fertility. The immediate and intermediate objectives as well as the ultimate goal are each subject to evaluation.

It is evident that, since evaluation is done in terms of objectives, the objectives of a program must be defined in some way so that success and failure can be measured. In other words, appropriate criteria must be established. This involves, among other things, specification of the nature of the objectives, the population that the program aims to reach, the time span within which change is to take place, and the magnitude of the desired effect (Suchman, 1967: 39-40).

The evaluation of family planning programs has not been a static process. With the development of programs over time and across countries, there has been increasing specification of the many interlocking activities involved in conducting large-scale programs. Global questions have been decomposed into component parts. Concurrently, more stringent criteria have developed for ascertaining success and failure at each stage and, in response, evaluation techniques have become more sophisticated, only to be subjected to increasingly stringent criteria.

This process can be seen by tracing the criteria for measuring demographic impact. Measures of contraceptive acceptance and length of use no longer suffice, and it is considered essential to measure the impact on fertility more directly. At the same time, it becomes the evaluator's task to incorporate the indirect effect of the program as well as the direct; to separate out the effect of substitution of program for non-program methods; and to take some account of non-program social and economic developments which may have a bearing on fertility. Evaluators are even asked to answer the ultimately unanswerable question of what would have happened in the absence of a program, as the yardstick against which to assess the impact of the program.

In response, the methods used for evaluation have been drawn from a number of disciplines and have ranged widely across the scientific repertoire within each discipline. One indication of the complexity of the questions posed is seen in the relation of evaluation research in family planning to basic research in demography. Though evaluative research is generally seen as residing in the area of application of existing knowledge rather than discovery of knowledge,

family planning research over time has added to demography in the refinement of various techniques, and through increased understanding of the process by which birth control affects fertility, to mention only some obvious examples.

IV. KEY PROBLEMS IN MEASURING THE EFFECT OF FAMILY PLANNING PROGRAMS ON FERTILITY

As will be seen in succeeding chapters, there are a number of difficult problems relative to the assessment of demographic impact that have not been solved to the satisfaction of interested scholars and administrators of family planning programs. These include the complexities of distinguishing direct and indirect program effects; determining what the secular trend of fertility would have been had the program not been undertaken; accounting for substitution effects; working out the future fertility of acceptors as compared with that of the general population; and interpreting evident changes in the fertility of acceptors. On another level, the evaluation of many programs is hindered by the lack of accurate data on fertility levels and on changes in fertility.

INDIRECT AND DIRECT PROGRAM EFFECTS

Because family planning programs combine a wide range of activities, including those designed to persuade all couples within a given area who are exposed to risks of conception to plan their families, there may be modification of values and behaviour even where this is not evident from acceptors and their performance. For example, couples may be influenced directly or indirectly by the program to practice birth control or to adopt a more effective method, but may obtain service through private channels. As yet, methods have not been found to ascertain to what extent the programs promote a diffusion of family planning in the general population, and some means of determining these indirect program effects is required to determine the full impact of family planning programs on fertility. A related problem is that some who have accepted a method under the program may drop out but transfer to private sources for service and supplies and continue to control fertility. A relevant question is whether post-program family planning on the part of such individuals should be credited to recruitment under the program and hence to births averted because of the program.

Assessment of the indirect effects of the program is further complicated by the non-program effects, i.e., declines of fertility not

causally attributable to the family planning program, which may be due to a number of factors, including changing demographic, economic and social conditions and cultural practices after commencement of the program.

DETERMINING FERTILITY CHANGE IN THE ABSENCE OF THE PROGRAM

Inevitably, the question arises as to what the secular trend of fertility would have been had there been no family planning program. The determinants of fertility and the mechanisms of fertility change are not sufficiently well understood to permit a definitive answer. If the setting is one in which fertility has been constant prior to introduction of a family planning program, the temptation may be to infer that all changes in fertility are due to the program. But to assess this with some assurance requires, among other things, information on the significant socio-economic and other relevant changes occurring during the same time-span, and what their impact would have been in the absence of a program. Efforts to cope with such problems have included the use of controlled experiments, matching studies and multivariate analyses of individual and areal data.

In addition to discovering the net impact of family planning programs upon fertility, there remains the compelling question of whether the programs initiate change or merely expedite already-existing trends. As the number and duration of programs increase, it should be possible to gain additional insight into this question.

SUBSTITUTION EFFECTS

A major problem in evaluating programs centers around what is commonly known as the substitution problem. This may take any of several forms: enrollment as an acceptor in the program by an individual who has been privately controlling fertility (this may or may not imply a change of method), and the shift of an individual from a less effective to a more effective method within the program are examples. A related problem is assessing the likelihood that new acceptors would have adopted fertility control in the absence of the program. Precise measurement of these effects has proved difficult, but such allocations are needed to determine the demographic impact of the program.

EXPECTED FERTILITY OF ACCEPTORS

Important in assessing the demographic impact of a program is the comparison between the observed and potential fertility of

acceptors. How best to estimate this potential has proved to be a difficult problem, given that acceptors are not a representative subsample of all couples. They tend to differ, for example, with respect to parity, prevalence of sterility, motivation to limit child-bearing, and experience of contraceptive practice. If the acceptors had not taken a method under the program, would their expected fertility have continued to be higher than that of a cross-section of the general population, and is the expected fertility of acceptors lessened by virtue of their comparatively high motivation?

SHORT-TERM AND LONG-TERM EFFECTS

When has a birth been averted? If the period of continued use without failure encompasses the amount of time that would elapse during the sequence of pregnancy, postpartum infecundity and waiting time while exposed to the risk of conception, can it be said that a birth has been averted? Or, as in the interpretation of period fertility data, is it equally plausible that a birth has merely been postponed? Can an assessment be made satisfactorily while the woman is still subject to the risk of child-bearing, or only when she is no longer susceptible, due to age, death of husband, divorce, etc.? What is the relevance of analysis of birth intervals, particularly the open interval? To what extent then may achievements in respect of contraceptive use amount only to spacing or short-term effects, and when do they reflect or forecast changes in completed family size?

V. SCOPE OF THIS VOLUME

The chapters that follow deal with the techniques employed and the major methodological issues that arise in evaluating the objectives of a family planning program. Though illustrations of the use of various techniques appear throughout the volume, the aim is not to assess the degree of success of family planning programs but to focus on the methods of measurement relevant to different objectives. The assumption here is that a better understanding of the potential and limitations of each technique will help make clear which aspects of program impact seem presently capable of evaluation. In this way, those responsible for evaluation will be aware where improvements in existing techniques are most needed, and at the same time, program administrators can become aware of criteria currently most amenable to evaluation. The inclusion of case studies from two countries (Chapters 11 and 12) and of various approaches

to target setting (Chapter 3) further highlight the degree of concordance between program objectives and the techniques available for their evaluation.

The chapters focusing on techniques address themselves to three levels of effect, with most attention given to the impact of a program on fertility, and less to the more immediate objectives. The division of the chapters according to the level of impact is as follows:

(1) On the most immediate level are techniques for assessing the extent to which the population makes use of the services offered ([1]). At issue here are determining the number and characteristics of acceptors, trends in acceptance, and rate of acceptance. The basic data are usually the "service statistics" generated from within the program. Chapter 2 discusses the concepts involved in such measurement, the methods of obtaining the data, and the manner in which such data can be utilized in evaluation. Retrospective surveys, discussed in Chapter 9, provide another means of obtaining data on the nature and frequency of contact with the program.

(2) A more intermediate level of evaluation focuses on how the use of a method from a program compares with the previous behavior of the acceptor, the length of time an acceptor makes use of contraception, and the action taken after termination of use. Answers to the first question can be obtained from certain types of service statistics. Retrospective surveys and follow-up studies of acceptors can develop data bearing on all three issues, particularly the last two. The utilization of such data in developing models for studying the use-effectiveness of contraceptives is described in Chapter 5.

(3) A number of techniques have been developed to measure more directly the impact of a program on some aspect of fertility and thereby attempt an evaluation of the more ultimate objectives. One line of work has been to estimate births averted, taking into account the nature of the contraceptive used, duration of use, and fertility prior to acceptance. These approaches require a detailed examination of the reproductive process and the role of contraception in reproduction. Chapter 4 provides the needed biostatistical background for appraising these techniques. Among the births averted techniques, one may distinguish attempts to measure the impact of contraception on period rates of fertility and those that trace the

([1]) Acceptance of services is not necessarily the most immediate effect sought by a program. Changes in knowledge and attitudes, for example, may be considered logically prior, but in terms of the focus of this volume on demographic impact, acceptance is the most immediate *behavioral* objective of a program.

impact on a cohort of acceptors. A further distinction can be made between empirical analyses based on the values of a specific population and simulation studies based on models of the reproductive process, in which a range of values for the parameters may be used and the effects of changes in different factors noted. Chapters 3, 6 and 10 illustrate these various births averted estimation techniques.

As distinct from births averted analysis, a number of procedures utilize a more holistic approach by attempting to appraise the demographic effect of a family planning program, taking into account social and economic developments which may also have an influence on aggregate fertility levels. The use of experimental design and matching studies, discussed in Chapter 7, and multivariate analysis of areal data, described in Chapter 8, illustrate two of these techniques. Data collected in a retrospective survey (Chapter 9) may permit one to make estimates of the role of acceptance vis-à-vis other determinants of fertility through the use of appropriate multivariate techniques applied to a model of fertility behavior. Such a survey can also be the basis of following a randomly chosen panel of couples for the purpose of studying their fertility in relation to their socio-economic characteristics and whether or not they become acceptors from the program. Surveys also provide a means of collecting data on many of the parameters used in births averted analysis.

In addition to these techniques, the case studies in Chapters 11 and 12 illustrate the measures of impact that project the pre-program trend of fertility and compare that with actual postprogram fertility, and others that correlate the fall in fertility with the trend in acceptors under the program.

Chapter 13 presents an overview of each of the techniques from the standpoint of its rationale, the sources of data, and the major conceptual and measurement problems associated with each approach. Some readers may find it helpful to start with Chapter 13 before pursuing the detailed explanations of each technique.

REFERENCES

Bean, L. L. and W. Seltzer

1968 "Couple years of protection and births averted: A methodological examination", Demography, Vol. 5, No. 2, Special Issue.

Berelson, B.

1969 "National family planning programs: where we stand" in S. J. Behrman, L. Corsa, Jr. and R. Freedman (eds.) Fertility and Family Planning: A World View, Ann Arbor, Michigan, p. 369.

1970 "The present state of family planning programs", Studies in Family Planning, No. 57 (September).

1974 "World population: status report 1974", Reports on Population/Family Planning, No. 15, January, The Population Council.

Berent, J.

1952 "Fertility and social mobility", Population Studies, Vol. 5, No. 3 (March).

Bogue, D. J.

1964 "The demographic breakthrough: from projection to control", Population Index, Vol. 30, No. 4 (October).

Caro, F. G. (ed.)

1971 Readings in Evaluation Research. New York, Russell Sage Foundation.

Chandrasekaran, C. and Moye W. Freymann

1965 "Evaluating community family planning programs", Public Health and Population Change, edited by Mindel C. Sheps and Jeanne Clare Ridley, University of Pittsburgh Press.

Coale, A. J.

1973 "The demographic transition", Volume I, Proceedings, International Population Conference, Liège, 1973. Liège, Belgium, International Union for the Scientific Study of Population.

Davis, K.

1963 "Scientist debate: will family planning solve the population problem?", Report of the Victor-Bostrom Fund for the International Planned Parenthood Federation. No. 10 (Fall).

Freedman, R.

1966 "Statement by the moderator", Proceedings of the World Population Conference, 1965, Volume I, Summary Report, United Nations.

Girard, A.

1951 "Mobilité sociale et dimension de la famille", Population, No. 1 (January-March).

Hauser, P. M.

1967 "Family planning and population problems - a book review article", Demography 4, Vol. 4, No. 1, pp. 397-414.

Keeny, S. M. (ed.)

1974 "East Asia review 1973", Studies in Family Planning, Vol. 5, No. 5 (May).

Lapham, R. and W. Parker Mauldin

1972 "National family planning programs: review and evaluation", Studies in Family Planning, Vol. 3 (March).

Mauldin, W. Parker

1967 "Measurement and evaluation of national family planning programs", Demography, Vol. 4, No. 1.

1968 "Births averted by family planning programs", Studies in Family Planning, No. 33, August, The Population Council.

Mauldin, W. Parker, N. Choucri, F. W. Notestein and M. Teitelbaum

1974 "A report on Bucharest", Studies in Family Planning, Vol. 5, No. 12, December, The Population Council.

Nortman, D.

1973 "Population and family planning programs: A factbook", Reports on Population/Family Planning, No. 2, fifth edition (September).

Notestein, F. W.

1967 "The population crises: reasons for hope", Foreign Affairs (October).

Reynolds, J.

1972 "Evaluation of family planning program performance: A critical review", Demography, Vol. 9, No. 1 (February).

Ridker, R.

1971 "Savings accounts for family planning, an illustration from the tea estates of India", Studies in Family Planning, Vol. 2, No. 7 (July).

Rogers, E. M.

1971 "Incentives in the diffusion of family planning innovations", Studies in Family Planning, Vol. 2, No. 12 (December).

Rossi, P. H. and W. Williams

1972 Evaluating Social Programs. New York, Seminar Press.

Seltzer, W.

1970 "Measurement of accomplishment: the evaluation of family planning efforts", Studies in Family Planning, No. 53 (May).

Suchman, E. A.

1967 Evaluative Research, Principles and Practices in Public Service and Social Action Programs, New York, Russell Sage Foundation.

United Nations

1964 Population Bulletin of the United Nations, No. 7, 1963, New York, Chapter VIII.

1972 Measures, Policies and Programs Affecting Fertility, with Particular Reference to National Family Planning Programs, New York.

Wang, C. M. and S. Y. Chen

1973 "Evaluating of the first year of the educational savings program in Taiwan", Studies in Family Planning, Vol. 4, No. 7 (July).

Chapter 2
ACCEPTOR DATA

by
C. Chandrasekaran

I. PURPOSE IN STUDYING ACCEPTORS

The use of the word "acceptor" in relationship to the provision of contraceptive services is relatively new and is associated with the development of large-scale family planning programs. Prior to the initiation of such programs, when family planning services were provided on a relatively small scale, mainly through private clinics, interest centered around the usefulness of contraceptives made available at these clinics. This was judged by qualities such as "acceptability", as gauged by the extent to which couples favored the use of the family planning method, and "effectiveness", which measured the protection it provided against pregnancy. Large-scale family planning programs require that goals be defined, that response to the program be assessed and that their ultimate effect be monitored and evaluated. In each of these tasks the use of "acceptor" and the associated concept of "acceptance" have been found to be of great value. In what follows we will discuss the problems that arise in developing the concept and definition of "acceptor", the methods used for obtaining data on acceptors and some of the uses to which data on acceptors can be put. In particular, we will deal with ways in which data relating to acceptors can be used for measuring program performance and progress. Several other uses that can be made of data on acceptors are discussed in detail in succeeding chapters; these are mentioned briefly in the concluding section of this chapter.

II. CONCEPTS AND DEFINITIONS RELATED TO ACCEPTORS

The expressions "acceptance" and "acceptors" have been used in a number of ways in family planning studies (United Nations, 1970, 56-59). Concepts of "acceptance" as used in the medical sense of

body-tolerance to a contraceptive or in the context of "group acceptance", to indicate the attitude of the social group to which an individual or couple belongs toward the practice of family planning in general or to the use of a contraceptive in particular, do not concern us here, although they are relevant to the organization of family planning programs.

An important problem which arises in defining "acceptance" or "acceptor" is the relationship which they bear to the beginning of the use of a family planning method. In the case of IUD insertions or sterilizations, acceptance is equivalent to the beginning of use. Conventional contraceptives and oral pills are more difficult to deal with.

An Expert Group (United Nations, 1969) considered the definitions to be attached to the terms "acceptor", "accepting" and "acceptance" in the context of an organized family planning program. In the view of this Group such terms would imply "the taking of service and/or advice from such a program, i.e., having an IUD insertion or a sterilization operation or receiving supplies (or advice on methods such as rhythm or coitus-interruptus) with the intent of using the method". This definition utilizes the concept of the intent to use the method, irrespective of whether the method was actually used. In the context in which the definition was stated, the taking of advice and/or services referred to the first time such advice or services were taken.

The comparative study of service statistics systems undertaken by ECAFE (United Nations, 1973) shows, however, that departures from this definition occur. In Malaysia, for instance, a person who accepts supplies (oral pills, condoms, etc.) for the first time when visited at a hospital, home or any place other than a family planning clinic is classified as a *provisional acceptor*. The person is motivated to return to the clinic to be registered as an *acceptor* and to obtain a re-supply of contraceptives.

A distinction is often made between "conception-regulating methods" and "birth-prevention methods" which has a bearing on the use of terms related to "acceptance". As "conception-regulating methods" aim at the reduction of the rate at which conceptions occur in the population, they do not include induced abortion, which is aimed at preventing a birth after conception has occurred. The view of the Expert Group referred to above was that the term "acceptance" should be restricted to contraceptive procedures whether physiological, mechanical or chemical, and including sterilization. Some have argued that the term "acceptance" should also include induced abortion.

For research and evaluation purposes a distinction has to be made between a "program acceptor" and a "method acceptor". A "method acceptor" is a first-time acceptor of a method from the program while a "program acceptor" is a first-time acceptor of any method from the program. In view of this distinction, it can be stated that an individual becomes a "program acceptor" at one point in time only, but he or she can become a "method acceptor" on several occasions.

Although both the concepts "program acceptor" and "method acceptor" stress the order of acceptance from the program, in practice, it is not always easy to determine whether an acceptance is in fact a first time acceptance or not. Information on this point is sometimes obtained while providing services or advice through the program. Since the number of "program acceptors" is an important measure of response to the program, it is sometimes necessary to estimate it through community surveys. However, close attention should be given to the hazard of underreporting of contraceptive use in surveys.

The stress given to "program acceptor" demonstrates that even in countries with large-scale national programs couples often adopt contraceptive methods without taking services or advice from the program. It is important that such non-program acceptances be considered in evaluation response to the program—first, because they reduce the potential number of "program acceptors", and secondly, because they may reflect a positive indirect effect of the program as, for example, through its public education activities. There has so far been no attempt to use an expression such as "indirect program acceptor". Attempts are now being made to establish a glossary of terms relevant to the evaluation of family planning programs. ([1])

Multiple acceptances by the same person from the program can involve two things:

(1) the receipt of the same method during several visits, such as obtaining additional supplies of oral pills,
or

(2) a change of method, such as a woman giving up the use of an IUD in favor of oral pills.

In the latter case the woman will be classified as an "IUD acceptor" as well as an "oral pill acceptor". Even when the same method is

([1]) By the Committee on Terminology of the National Family Planning Forum. See report in *Perspectives*, volume 4, No. 5, July 1972, Planned Parenthood Federation of America, Inc., New York.

accepted on two or more occasions, there could be a break in its use, as for instance in the case of a reinsertion of an IUD following the expulsion of one inserted earlier. In this example, the woman may in some countries be considered an IUD acceptor, although for the purpose of evaluation, the break in continuity of use will have to be taken into account (Tietze and Lewit, 1973). The Philippine and Indonesian programs instruct staff to count a woman as a "first acceptor" only once, regardless of resupply visits or switching methods. In the Korean program, on the other hand, the same woman can be counted as a first acceptor, at different times, of the loop and the pill, and her husband of vasectomy.

As a program continues to operate for a long period, it is difficult to keep track of the changes in contraceptive practice of any particular individual, especially when pregnancies, planned or unplanned, intervene. The expression *old acceptor* is used in the programs of Hong Kong, Indonesia, Nepal, Sri Lanka and Thailand to differentiate *first acceptors* from those who had already received advice/service from the program.

Acceptance, according to the definition given earlier in this chapter, implies the taking of services and/or advice (for the first time) from a program. As such, it is the initial step in the beginning of use of a method given or advised by a program. In the case of most contraceptive methods, possibly with the exception of the IUD (and, of course, sterilization), acceptance does not necessarily signify use of the method, though it usually implies an intention of use (Chandrasekaran and Karkal, 1972). The extent of use after acceptance from the program will be dealt with in Chapter 5. Relevant, however, to the evaluation of the effect of acceptance is also the status of the acceptor with respect to the use of one or more methods of contraception prior to acceptance. The following two alternative schemes have been proposed for classifying couples (or clients) according to use or non-use of contraceptives:

Scheme 1	*Scheme 2*
(a) Never users	(a) Current users
(b) Ever users	(b) Current non-users
(i) Current users	(i) Never users
(ii) Past users	(ii) Past users

Note that each of the entries involves one or more cells from a simple 2 by 2 table that crosses use now—Yes/No, against use in past—Yes/No.

III. SOURCES AND PROCEDURES FOR OBTAINING DATA ON ACCEPTORS

The major procedures for securing data on acceptors are through the use of

(1) the individual client record, sometimes called the clinic record or the case card,
(2) the coupon system,
(3) the registers, or
(4) surveys, and the processing of information obtained from these sources.

The individual client record is the logical extension of the system normally used in advanced medical or public health practice, and as a rule treats the person or the couple seeking family planning advice as requiring the same amount of care and attention as those seeking medical advice. In the early stages of the family planning movement, when services were being provided through private family planning clinics, these clinics made use of the individual client record. Even in large-scale family planning programs, the individual client record is the method of data collection that is most widely used. The use of the coupon was first introduced in the family planning program in Taiwan and was also quickly adopted by the Korean program. The coupon is simpler in its format than most of the individual client records now in use, and as will be discussed below, was developed partly to suit the monetary payment system on which the Taiwan program was based and partly to facilitate quick processing of data on acceptors. The use of registers, like the use of the individual client record, is the extension of a practice which has been followed in the provision of medical and health services in some countries. The three methods—the individual client record, the coupon and the register—aim primarily at providing information on "program acceptors" and "method acceptors" as defined above; however, there are difficulties in avoiding multiple counts of the same individual through the use of these methods when services are provided through a large number of clinics. The difficulty of recording a continuous history of contraceptive use by acceptors as a program becomes long-standing has also been pointed out previously. The survey procedure—the fourth method listed above for obtaining data on acceptors—can retrospectively provide information on acceptances through the program as well as from outside the program although, as pointed out earlier, it should be recognized that surveys are likely to underestimate

contraceptive use. Since the survey procedure is adequately dealt with in Chapter 9, the present discussion will center around the use of the individual client record and the coupon, with a brief reference to the use of the register.

INDIVIDUAL CLIENT RECORD

The purpose of the individual client record is primarily to enable the clinic personnel to provide the best advice to the individual client. The record form when properly drawn up and carefully maintained can also provide information for guiding administrative decisions on program activities and policies. The individual client record is being used in almost all family planning programs. Although details differ, the content of the forms used in the different countries is broadly similar. It can be grouped under the following headings:

(i) Identifying particulars.
(ii) Social and economic characteristics having a bearing on the methods to be advised.
(iii) History of pregnancies and their outcome and number of children living.
(iv) Past and current practice of contraceptive methods.
(v) Reasons for seeking family planning advice.
(vi) Results of medical examination conducted at the clinic.

Much of the information obtained under the above headings, as indicated in later chapters, can also guide the administration in evaluating and improving the program. Some information is also obtained on the individual client record specifically for the purpose of assisting the administrator. An example is the source of referral of the "acceptor" to the clinic, which can assist in gauging the influence of different educational activities of the family planning program in drawing clients to the clinic.

A regional seminar held in Bangkok (United Nations, 1970) listed the following core items of a non-medical nature for inclusion in the individual client record.

(a) Name and address of the clinic/hospital/other center for family planning service
(b) Identification number of acceptor
(c) Name of acceptor
(d) Name of spouse of acceptor
(e) Complete address of acceptor
(f) Date of first consultation at this clinic

(g) Date of birth of acceptor or present age in completed years
(h) Age of spouse
(i) Education of acceptor in completed number of years of schooling
(j) Education of spouse in completed number of years of schooling
(k) Monthly family income in broad categories
(l) Number of children ever born, by sex
(m) Number of living children, by sex
(n) (i) Desire to have more children in future
    (ii) Preferred interval after which next child is desired
(o) Date and nature of termination of last pregnancy of female spouse
(p) (i) Ever-use of family planning method in the past
    (ii) Methods used in the past
    (iii) Last method used
(q) Date of acceptance of family planning method from this clinic
(r) Referral agency/agent.

The seminar also recommended the following items as optional: race, religion and marital status. The recommendations of the seminar pertaining to medical aspects also appear in its report. The seminar was of the view that the list of items on the client record form should be as short as possible, in view of the consistent under-use of the record for analytical purposes in many countries. A typical client report which can be adapted to different circumstances is also presented, without the required medical data, in *A Handbook for Service Statistics in Family Planning Programs* (Ross, 1968).

ECAFE (1973) made a study of the information obtained in the individual client cards of ten countries (Fiji, Hong Kong, India, Indonesia, Malaysia, Nepal, the Philippines, Singapore, Sri Lanka and Thailand). Age of acceptor, number of living children and open interval were found in the client cards of all these countries. Other demographic characteristics on which information was obtained included the number of living children by sex (all except Sri Lanka), and the number of live born (all except Sri Lanka and Thailand). The purpose for using contraception (spacing vs. limitation) was sought in all programs except Fiji and India. The programs in Sri Lanka, Hong Kong, Malaysia and Singapore recorded information on the duration of marriage. Relatively less uniformity was found in the information obtained on social and economic characteristics.

As discussed in Chapter 6 two characteristics of the acceptor are found important in the calculation of the number of births averted by the family planning program. One relates to contraceptive practice at the time when a person becomes an acceptor, from which it would

be possible to determine whether the program was substituting for a method already being used by the acceptor. It should be possible to obtain the relevant information on the individual client record. The other characteristic is the "menstruating status", which is especially important in post-partum programs. It is, of course, possible to inquire about the last menstrual period on the individual client card and this is usually done; however, if the woman is in post-partum amenorrhea, the date of resumption of menstruation has to be obtained during follow-up visits or by requesting the woman to inform the clinic of this date.

An individual client record normally has items and space in which relevant data on revisits of the client to the clinic are recorded. Such data include date of revisit, use of method, pregnancy status, complications, if any (especially as related to the use of the method), and new method prescribed. If the contact between the client and the clinic is close, the individual clinic record can give a continuous picture of the experience of the client from the time of first acceptance of a family planning method from the clinic, including continuity of use of the method, change over to other methods, occurrence of pregnancies, nature of termination, etc. In practice, however, the individual clinic card cannot be depended upon to give such continuous history and for several evaluative purposes it is necessary to obtain additional information through surveys.

THE COUPON SYSTEM

The coupon system can be exemplified by the one used in Taiwan for IUDs. Each coupon consists of three parts with the same serial number printed on all three. Booklets containing twenty coupons are given to the field worker who visits the prospective acceptor in her home. Part I of the coupon remains in the booklet from which parts II and III are detached and given to the potential acceptor. Part I contains the name of the woman visited, her address, and the date of issue of the coupon. This information is important to the field-worker for follow-up purposes. Generally, a coupon acceptor is followed up six to eight weeks after the date of acceptance to find out whether she has gone to the doctor for an insertion.

During the home visit, the fieldworker enters in part II such particulars as the recipient's name, age, address, level of education, number of living children by sex and date of issue. The fieldworker also designates herself in part II as the referral agent to provide feedback for supervisors who will evaluate her performance. The physician in the clinic adds such information as the clinic address. After

inserting the IUD he sends part II, which contains most of the necessary information, to program headquarters for claiming payment and for processing. Part III is retained by the inserting physician for purposes of payment. The coupon system has been extended to cover sterilization operations and pill acceptance in Taiwan.

The strength of the coupon system, as well as the difficulties that arise in its operation, were discussed by an Expert Group (United Nations, 1969), which made the following recommendations:

(a) The coupon system, as outlined above, is a simple and efficient procedure for the collection of data needed by administrators and evaluators, and should be tried in countries where one-time methods such as IUDs and sterilization are publicized, especially with monetary incentives attached to the scheme.
(b) It is essential to have a central evaluation team of highly qualified personnel to guide and evaluate the coupon system.
(c) Consideration should be given to designing a coupon with four parts: part I to be retained by the fieldworker, part II by the doctor, part III to be forwarded for evaluation and part IV to be sent as a claim for payment.
(d) The part of the coupon which is retained by the clinic or doctor should be elaborated to the greatest extent possible to serve the additional purpose of a clinic card.

REGISTERS

A register is maintained in some clinics to provide information on the persons who visit each day. When such registers are maintained in addition to the individual clinic record, they do not provide any additional information. Where no individual clinic record is kept, however, the register, bearing as it does a line entry for each woman who visited, can yield simple daily data on contraceptives provided, such as the number of IUDs inserted or pill cycles distributed.

IV. USE OF ACCEPTOR DATA FOR EVALUATING PROGRAM PERFORMANCE

Acceptor data obtained through service statistics—individual client record, coupon system or registers—have to be processed and analyzed if they are to assist the program administrators and others concerned about program achievements. Some examples of rather complex investigations that can be undertaken using data on acceptors will be mentioned in subsequent chapters, especially chapters 5,

6 and 8. The discussion in this and the following sections will deal mainly with what can be learned from routine analyses of acceptor data.

Mention has already been made of the chief difficulty that arises in obtaining accurate data on "program acceptor" and "method acceptor", the danger of multiple counts if the same person takes advice from two or more clinics served by the program. Information on previous attendance at other clinics run by the program, if obtained routinely on the individual clinic record, can assist in minimizing such duplication in "program acceptor" counts. Similar data on past methods accepted will help avoid duplication in the case of "method acceptors". It should of course be recognized that the information obtained from the acceptors on age, number of living children, etc. is likely to contain errors, just as in the case of other data obtained through enquiries, especially in developing countries. The adjustments to be made are beyond the scope of the present discussion.

NUMBER OF ACCEPTORS

The monthly number of program acceptors or method acceptors, as shown in section V of this chapter, can give useful guides of program progress. In countries where targets for acceptors are set, the achievements of the program can be assessed by comparing the number of acceptors directly with the target. This procedure was frequently used in Taiwan and Korea, the two countries which first

Table 2.1. *IUD acceptors by county/city: January 1967*

County/city	*Monthly acceptors*	*Monthly target*	*Percentage of target*	*Cumulative acceptors*	*Percentage of cumulative acceptors to married women 20-44 years*
1. Keelung city	157	214	73.4	6,690	19.0
2. Taipei city	667	1,200	55.6	20,167	14.0
3. Taichung city	251	204	123.0	11,736	24.5
...	...	...	...	...	...
...	...	...	...	...	...
22. Penghu county	16	144	11.1	781	6.0
Grand total	7,901	10,833	72.9	268,646	16.9

Source: Monthly Report of Taiwan Family Planning Program, January 1967, Taiwan Population Studies Center, Taichung.

began setting targets. The ratio of the number of acceptors to the number set as a target, expressed as a percentage, can serve as one of the achievement indices. In Korea, for example, the target of IUD insertions for 1965 was 200,000 while the number of actual insertions was 223,000, giving an achievement index of 116.5 per cent. An adaptation of the routine statement prepared in the Taiwan program is given in Table 2.1.

When targets are set for acceptance of different methods such as IUDs, sterilizations or traditional contraceptives, there is an interest in studying the achievement of the entire program including all the methods provided. No such index has so far been developed using only acceptor data. A composite index of the output of a program, including new acceptors and returning clients, in which different methods are offered was proposed by Wishik and Chen (1973) in connection with the Pakistan program, using data on "applications" such as number of IUDs inserted and number of condoms and contraceptive pills distributed during a year. According to Wishik, this index, called the "couple-year of protection" (CYP) can be used as an achievement index. It is based on the number of "applications" of each contraceptive weighted according to the protection it affords, regardless of when that protection is in effect. Conversion factors must be developed specifically for each place. For example in Pakistan, the weight or "conversion factor" for a condom is assumed to be .0083 on the basis that it requires 120 condoms to protect a couple using this method for the period of one year. The conversion factor is assumed to be 2.5 CYP for an IUD and 0.25 CYP for an injection. If during a year a program administered N_1 applications of contraceptive i with conversion factor f_1, the CYP achievement during the year is given by $\Sigma N_1 f_1$, where Σ represents summation over all the contraceptives supplied.

Data on acceptors become more meaningful when converted into a *rate*. This is true whether one deals with the number or the characteristics of acceptors. Confining attention—for the present—to the number of acceptors, the logical way of converting it into a rate for measuring coverage by the program is to relate it to a suitable base. Such a base should be the number of eligible women from whom the acceptors are drawn into the program. Some couples are temporarily or permanently ineligible for accepting advice from a program. If the husband or wife has been permanently sterilized, the couple has no further need to accept a family planning method. It is also apparent that a couple has no need to use a method while the wife is pregnant. Couples who are currently using a family planning method and are well-satisfied with it are less likely to look for a new

method and therefore have less chance of accepting a method from the program than couples who are not using any method or are not satisfied with the methods they are using. Using data from the Taichung family planning study. Takeshita and Freedman (1967) have illustrated the use of the "eligibility" concept on a progressively narrowing basis starting with all married women of a given age-range and successively excluding (a) the sterilized or sterile, (b) those who wanted more children and were not using any contraception, (c) those who were pregnant, and (d) those who were using some type of contraception when they were first visited by program staff. They found that relating acceptances to varying eligibility bases resulted in a series of alternative rates sensitive to different aspects of program objectives.

In the same study, Takeshita and Freedman showed that by using the eligibility concept it was possible to decompose the acceptance rate (the rate based on all women) to show "that it is a function not only of

(1) the rate of acceptance among women who are defined as eligible by a given criterion but of
(2) the proportion of women who are so eligible and of
(3) the spill-over ratio which measures the extent to which there is acceptance among those who are not eligible by the criterion used".

The eligibility criterion will vary with the type of program and the emphasis placed on certain contraceptive methods. In a postpartum program, the eligible women are those who are pregnant. In a program with an accent on sterilization, age and parity characteristics are relevant. In a comprehensive program, the number of eligible persons changes with time. In fact, as a program advances, the number of acceptors drawn into it increases, diminishing the reservoir of additional persons who can become acceptors. This reservoir is also affected by the entry of newly-married women and the exit of married women because of death, widowhood or menopause. For the most part, workers in the field of family planning evaluation have been satisfied with the use of simple though imprecise indices, mainly because of the difficulty of obtaining more adequate data.

An obvious group to which the number of acceptors can be related is the number of currently married women in the reproductive age-group. This procedure has been used routinely in Taiwan. In some cultures where reproduction occurs without formal marriage, the base to be used should appropriately be all women in the reproductive age-group. To illustrate the former procedures, the average

number of IUD insertions per year during the three-year period 1965-1967 in Korea was about 300,000 which formed 8.2 per cent of married women in the age group 20-44 years. In Taiwan the number of IUD insertions per year during the same period averaged 110,000 amounting to 6.9 per cent of married women in the age-range. In Singapore, about 30,000 or roughly 14 per cent of married women in the age-group 20-44 years became acceptors each year during the three-year period 1966-1968. Of them, the large majority accepted the pill which was the method favored by the program (Chandrasekaran, 1969). Even this relatively simple method for measuring acceptance cannot easily be applied in practice because of the difficulty of obtaining data on the number of currently married women except in census years.

ACCEPTOR CHARACTERISTICS

Ross *et al.* (1972) have listed the following four measures which they used to study the demographic characteristics of acceptors in a number of national family planning programs. ([2])

1. *THE ACCEPTANCE RATE.* This gives the annual number of acceptors per 100 married women, by age and by other characteristics as data permit. It is calculated by dividing the number of acceptors in a sub-group by the number of married women in that sub-group, and multiplying by 100.

2. *THE ACCEPTANCE INDEX.* When the numerators and denominators for the calculation of the acceptance rate are not available, but the percentage distribution of acceptors by a characteristic and the percentage distribution of married women by the same characteristic are available, an acceptance index can be calculated. This index shows under- or over-representation of acceptors in categories of a characteristic such as age or number of living children, compared with the distribution of married women in general. The index is calculated by dividing the percentage of acceptors in a specific category by the percentage of married women in that category, and multiplying the ratio by 100.

3. *THE PERCENTAGE DISTRIBUTION OF ACCEPTORS* by various characteristics is the third measure. As the authors have pointed out,

([2]) All four had been previously used by one or more workers in the study of acceptors, although they were not designated by specific titles.

this index "lacks the advantage of acceptance rates and indices, namely, controlling for differences in the denominators from which acceptors come through time or between populations".

4. *THE MEDIAN* can be considered as a useful statistic for comparing different populations by a specific characteristic. Like the *percentage distribution of acceptors*, from which it can be derived, the median suffers from the deficiency that it does not control for differences in the base populations belonging to different categories of the characteristic.

Data obtained from the individual client record or the coupon can be used directly for working out the *percentage distribution of acceptors* or *the median* for a specified characteristic found in the records, such as age of acceptors. Further, it is possible to use these data to provide cross-classification of two or more characteristics found on the individual client record or the coupon. One such example, which will be discussed later, is the classification of acceptors by "method accepted" and some other characteristics—age of acceptor, number of living children, etc.

Although, as indicated above, the *acceptance rate* or the *acceptance index* carries greater precision than the other two indices mentioned, it is not usually possible to calculate them other than for a few characteristics, such as age of acceptor, number of living children or educational level, for which censuses can provide rough data for the base population. Periodic community surveys can be used to provide base data, but such surveys have been carried out in only a few countries, notably Korea.

It is interesting to illustrate the use of the above indices by applying them to selected acceptor characteristics recorded in some of the programs. For the most part the material provided below is derived from other publications and the reader is referred to these sources for the limitations of the data used and any assumptions that were made.

5. *AGE.* The age of the acceptor is one of the few characteristics for which attempts to calculate the acceptance rate have been made. Results of a comparative study made by Chandrasekaran for Taiwan, Korea, Singapore and Hong Kong are presented in Table 2.2. The study showed that the age pattern of acceptance around 1967 differed among the countries. In Singapore, where the general response was very high, the highest age-specific acceptance rate was recorded for women in the age group 20-24 and the rate declined progressively with advancing age. A similar trend was observed in Hong Kong. In

Taiwan and Korea the patterns were different. In Korea, the highest rates were recorded for age-groups 30-34 and 35-39, while in Taiwan the highest rate was found in the age group 30-34 years.

Table 2.2. *Number of annual acceptors per 1,000 married women, by age group, selected countries*

Age group	*Taiwan (1967)*	*Korea (1966)*	*Singapore (1967)*	*Hong Kong (1967-1968)*
20-24	54	27	260	93
25-29	87	79	191	59
30-34	95	146	126	48
35-39	71	148	78	30
40-44	42	87	55 (a)	17 (a)

(a) Calculations based on data of acceptors aged 40 and over in the numerator and married women aged 20-44 in the denominator.

Note: Figures for Taiwan and Korea refer only to first loop insertions. The figures for Taiwan are given in Chow, L.P. and Chen, H.C. "Evaluation of the family planning program in Taiwan", *Journal of the Formosan Medical Association*, Volume 67, No. 7, July 1968, Table 4. Figures for the other three have been calculated from data on acceptances published or obtained from the programs. Singapore figures are based on the data on acceptance for the six months July to December 1967 but are shown on an annual basis. Data on base populations used in the calculations were derived from census or survey publications. Figures given for Korea, Singapore and Hong Kong should be considered approximate.

Source: Chandrasekaran C., "Family Limitation in South and South-East Asia, International Population Conference, London, 1969; The International Union for the Scientific Study of Population, p. 1246.

More recently Ross *et al.* (1972) have presented acceptance rates by age for Hong Kong, the Republic of Korea and Taiwan, and acceptance indices by age for a larger number of countries. The authors came to the following conclusion with respect to variations in acceptance by age, taking into account all methods:

> "Acceptance rates and indices increase with age to 30-34 years for most of the countries, after which they start to decline. Most acceptors are therefore between 25 and 34 years old although simultaneously appreciable proportions are younger than 25 or older than 34. Acceptors in Hong Kong and the Philippines tend to be more heavily represented in the younger age categories of 20-29 years; in South Korea they are more heavily represented at 30-39."

The pattern in Singapore, which was not taken into account by Ross *et al.*, as mentioned earlier, shows a declining acceptance as age advances.(3)

(3) The acceptance rate using all women in the age-group as base shows an increase up to age-group 25-29 and a decline thereafter for the years 1968 and 1969 (Singapore Family Planning and Population Board, 1969).

Table 2.3. *Distribution of acceptors, and acceptance rates of loop, by number of living children, 1964-1967*

Number of living children	*Distribution of acceptors (per 1,000)*				*Acceptance rate (a)*				*Ratio 1967-1964*
	1964	*1965*	*1966*	*1967*	*1964*	*1965*	*1966*	*1967*	
None	1	3	2	1	1	3	3	2	314
One	18	23	30	30	5	14	21	22	423
Two	92	108	126	134	18	44	57	67	371
Three	207	214	229	243	37	80	95	91	244
Four	256	249	248	250	42	84	93	92	220
Five or more	426	403	365	342	38	75	75	81	212
Total	1,000	1,000	1,000	1,000	30	63	70	74	243
Actual number of total acceptors	46,600	99,253	111,242	121,053	46,600	99,253	111,242	121,053	—

(a) Rate per 1,000 married women in the respective group in Taiwan.

Source: Chow, L. P., "Evaluation of the family planning program in Taiwan". Republic of China. Reprinted from *The Journal of the Formosan Medical Association*, vol. 67, No. 7, July 28, 1968.

Several factors are likely to influence the age pattern of acceptance. These include the family planning methods offered by the program, the length of time for which the program has been in operation, the extent to which contraceptive practice prevailed before the start of the program, or was resulting from non-program influences and, above all, the motivation in the community for practising family planning. Such motivation will also express itself in other characteristics of acceptors such as number of living children or educational level, which are likely to be related to the age distribution of acceptors.

NUMBER OF LIVING CHILDREN. Taiwan's is perhaps the only program in which *acceptance rate* by the number of living children of women having loop insertions was worked out annually. The annual rates obtained in this program for the years 1964 to 1967 are presented in Table 2.3. The increase in the rate when the number of living children advances from none to four is well brought out by the figures given in this table. Ross *et al.* (1971) have provided data on the acceptance index for five programs. Their conclusion is that taking into account all methods offered by the program, "Acceptance rates and indices rise with the number of living children, at least to about five to six."

It is interesting to compare the trend in *acceptance rate* by number of living children as given in Table 2.3 with the proportion of acceptors (per 1000) by number of living children as given in the same table. The latter is equivalent to the index *percentage distribution of acceptors* referred to above. The difference in the relative values for five or more living children as compared with the respective values for four living children for the two indices emphasizes the danger of using *the percentage distribution of acceptors* as a substitute for *acceptance rate.*

It is rare that acceptance has been studied by considering at the same time both age and number of living children. While it is easy to classify acceptors through a cross-tabulation of age and number of living children, a similar tabulation for the appropriate population to be used as the base for the calculation of acceptance rates is not usually available. Acceptance rates for IUDs by age and number of living children have been calculated for the Taiwan program (Chow, 1968). Wolfers (1970) worked out such rates for the different ethnic groups in the Singapore IUD post-partum program. Such cross-classifications indicate that for a fixed number of living children acceptance rates usually decline with increasing age. In the Bombay study (Chandrasekaran and Kuder, 1961) the acceptance rate declined with age when the number of children born alive was held

constant. As an extension of this type of analysis the variation in acceptance rates can be studied through a cross-tabulation of age with the number of living male or female children. Wolfers found a marked increase in the acceptance rate between women with no living male children and those with one living male child, especially among the Chinese.

OTHER CHARACTERISTICS. Ross *et al.* (1972) have provided a summary of their findings with respect to the following characteristics: open interval (interval from last birth to acceptance), previous use of contraception, residence (rural/urban) and education. The findings are based on the experience of very few countries and are limited to the index *percentage distribution of acceptors.* Excerpts from the findings are reproduced below:

> "*Open interval.* There is wide variation in the interval between last birth and acceptance. The median interval ranges from four to twenty-nine months; from 6 per cent to one-third of women accept a method of contraception within six months (of their last birth).... The interval is shortest in Singapore which has a postpartum oriented family planning program and where the pill has been for some time the most widely accepted method. The interval is longer for Thailand and Taiwan and for Pakistan in 1969; the longest intervals are found in India and South Korea.
>
> *Previous use of contraception.* The percentage of couples who had not used contraception before acceptance is high, ranging from 63 to 89 per cent; this suggests that the program is serving many couples who might not otherwise have practised contraception.
>
> *Residence.* In general, urban areas have higher acceptance indices than do rural areas. The two exceptions are South Korea and Taiwan (which) have well-developed rural family planning programs.
>
> *Education.* The majority of acceptors in family planning programs have less than a secondary school education. A substantial proportion of acceptors are illiterate, but there is wide variation among countries."

METHODS ACCEPTED. The analysis of methods accepted from the program is usually undertaken by using the information found on the individual client record or its equivalent. Interest normally centers around assessing:

(1) the choice of the different methods offered by the program, and

(2) the relationship of various social and economic characteristics to the acceptance of the different methods.

For both purposes the index most widely used is the *percentage distribution of acceptors.* In some quantitative instances the *median*

values of the social or economic characteristic for the different methods are compared. An application is illustrated below.

OVERALL ACCEPTANCE OF DIFFERENT FAMILY PLANNING METHODS

A typical representation is that provided in the *Fourth Annual Report* (1969) of the Singapore Family Planning and Population Board which is reproduced as Table 2.4. Since 1966, the contraceptive pill is the most widely adopted method, followed by the condom.

Table 2.4. *Singapore: Acceptance of family planning methods by percentages, 1964-1969*

	1964	*1965*	*1966*	*1967*	*1968*	*1969*
			(as percentages)			
Intra-uterine device	—	11	10	1	11 (*a*)	4 (*a*)
Contraceptive pill	20	33	46	61	55	51
Condom	46	36	27	31	29	39
Other methods	34	20	17	7	5	6 (*b*)
Number of patients	9,339 (SFPA)	9,845 (SFPA)	30,410 (FPPB)	30,935 (FPPB)	35,338 (FPPB)	35,643 (FPPB)

(*a*) Includes acceptors with IUD *In situ* from Government-University Research Project and FPA.
(*b*) Includes 1 per cent injections.
Note: SFPA: Singapore Family Planning Association. FPPB: Family Planning and Population Board.
Source: Fourth Annual Report, 1969, Singapore Family Planning and Population Board.

Table 2.5. *Hong Kong: Percentage distribution of new patients by method accepted* (*a*)

	1968	*1969*	*1970*
Oral pill	45	60	68
Intra-uterine device	21	13	9
Condom	14	11	8
Injectables	4	3	5
Sterilization applied for	3	3	3
Diaphragm, jelly etc.	1	1	2
Others and non-users	12	9	5
Total	100	100	100

(*a*) The new acceptors numbered 26,588, 30,886, and 30,470, respectively, in 1968, 1969 and 1970.
Source: Twentieth Annual Report, 1970-71. The Family Planning Association of Hong Kong. pp. 12-13.

IUDs, injections and female sterilization account for less than 10 per cent of total first acceptances. Male sterilization has not been an important method in Singapore, only three having occurred among a total acceptance of about 35,000. The percentage distribution of new patients by method accepted in Hong Kong during the years 1968, 1969 and 1970, and given in Table 2.5, again shows the preference for oral pills. However, in Hong Kong, unlike Singapore, a larger number of acceptors chose the IUD rather than the condom.

Dorothy Nortman (1972) has compiled data giving information on the number of acceptors in several countries by family planning method used. There is wide variation in the patterns of acceptance. In Chile during 1969, for example, 71 per cent accepted IUDs, 22 per cent oral pills and the remaining 7 per cent other methods, excluding sterilization, while in Colombia during 1971, 51 per cent accepted IUDs, 46 per cent pills and the remaining 3 per cent other methods excluding sterilization. In Costa Rica, Egypt and Kenya the majority of acceptors take oral pills and in some countries, such as Iran and Western Malaysia, those accepting oral pills exceed 90 per cent. Sterilization, especially vasectomy, is widely accepted in India, 57.0 per cent of the 13.6 million acceptors during 1965-71 having accepted this method.

It should be recognized that the percentage figures worked out for different countries cannot be used as a measure of the extent of preference for one or another method in these countries; for all the methods are not made equally available in the programs of the different countries. India, for instance, does not provide oral pills in the national program, while the programs in Iran and Western Malaysia rely almost entirely on pill distribution.

ACCEPTANCE OF METHODS BY SOCIAL AND ECONOMIC CHARACTERISTICS

The association between a method accepted and a social, demographic or economic characteristic is displayed by working out a two-way table between the method accepted and the categories for which information on the characteristic is available. When the characteristic is of the qualitative type, such as "ethnic group", the percentage of acceptors belonging to each ethnic group who accepted the different methods provides the basis for understanding the relationship between the method accepted and ethnic group. Table 2.6 provides an illustration of this type of representation drawn from the experience of Singapore. The data given in the table show that

Table 2.6. *Singapore: New acceptors by ethnic group and method, prescribed 1968 and 1969*

Method prescribed	Percentage of population by ethnic group and method, 1968 and 1969									
	Chinese (percentage)		Malay (percentage)		Indians and Pakistanis (percentage)		Others (percentage)		Total percentage	
	1968	1969	1968	1969	1968	1969	1968	1969	1968	1969
porary	24	22	22	20	23	23	14	18	23	21
a-uterine device	11	3	6	1	16	5	9	10	10	3
	49	49	60	55	44	46	66	61	51	51
doms	14	23	10	21	12	24	4	3	13	22
ers (a)	2	3	3	3	4	2	7	7	3	3
l	100	100	101 (b)	100	99 (b)	100	100	99 (b)	100	100

(a) Including spermicides and injections.
(b) Do not add to 100 because of rounding.
Source: *Fourth Annual Report, 1969.* Singapore Family Planning and Population Board.

the Malays accepted the pill to a greater extent than the Chinese or the Indians and Pakistanis, and the IUD to a lesser extent. Where the characteristic to be studied is of the quantitative type, such as age of acceptor, number of children born to acceptor or the number of years of schooling she has undergone, the *median* or *average* value of the characteristic for each of the methods accepted can be provided. The *median* is preferred to the *average*, although in general the conclusions reached will not differ greatly by using either index. The average age, number of children born and years of schooling of the acceptors in the Hong Kong program for the methods accepted during the years 1969 and 1970 are given in Table 2.7. Explaining these data, the report from which they have been taken (Hong Kong, 1971) states:

> "... oral pills were chosen by those women who were young, low in parity and better educated. Injectables were in general accepted most readily by older women of high parity and least well educated while IUDs and conventional methods were adopted by women in median categories."

Ross *et al.* (1972), after a review of the data obtained in a number of programs, have summarized their findings on the age distribution and number of living children of acceptors of different methods as follows:

Table 2.7. *Hong Kong: Average age, number of children and years of schooling of acceptors of various methods*

Method	Age		No. of children		Years of schooling	
	1969	1970	1969	1970	1969	1970
Intra-uterine device	32.4	32.6	3.7	3.8	4.0	4.3
Oral pills	27.7	27.2	2.3	2.2	4.7	4.9
Injection	35.4	35.1	4.7	4.9	3.2	3.0
Other methods	32.3	32.7	3.5	3.3	3.7	4.3
All acceptors	29.5	29.0	2.8	2.7	4.4	4.6

Source: Twentieth Annual Report, 1970-1971. The Family Planning Association of Hong Kong.

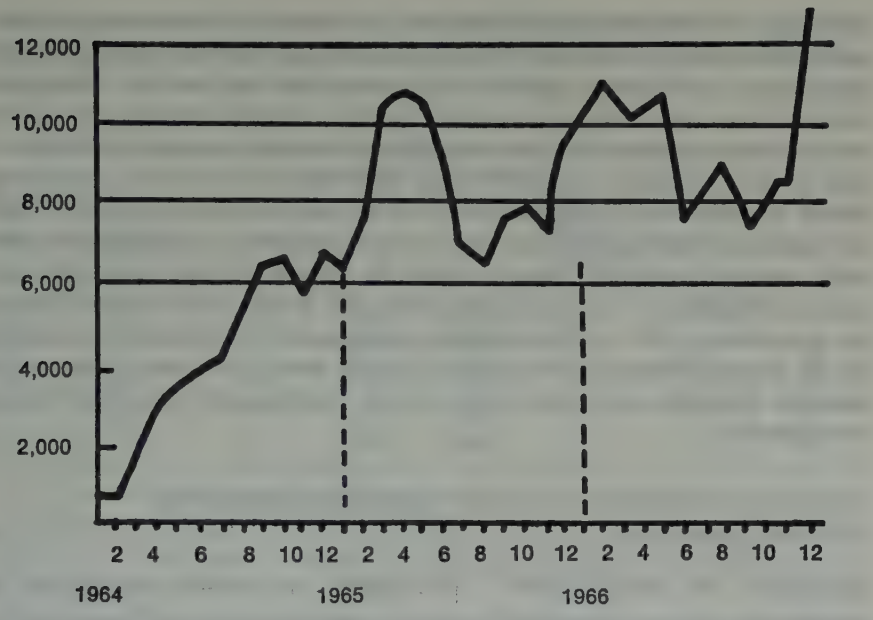
"*Age.* Most acceptors of IUDs are between 25 and 34 years old, but substantial proportions are younger than 25 and older than 34. The median age of IUD acceptors ranges from 28.8 to 33.8; in most national programs the median age is about 30 years. Acceptors of pills are somewhat younger than IUD acceptors. The median age in most countries is about 29. Most acceptors of vasectomy are over 35 years old. Wives of vasectomy acceptors, though considerably younger, are still older than women who accept IUDs. The median age of wives ranges from 30.9 to 36.2 years. The former figure refers to India, where vasectomy is the most popular method.

"*Number of living children.* In the case of the IUD as well as the pill, roughly 27 to 51 per cent of acceptors have three or four living children. For vasectomy, 35 to 49 per cent of acceptors have three or four living children. As is to be expected, a higher proportion of the acceptors of IUDs and pills have fewer than three living children than the number for wives of men who undergo vasectomy operations; and similarly the latter group has a higher proportion with five or more living children.

"It should, however, be noted that there are considerable variations in the characteristics of acceptors choosing the same method in different programs. The median number of living children for persons undergoing vasectomy at Ernakulam in Kerala State was 3.6 which compares favorably with the median number of children born to IUD acceptors in several countries. Also, the median number of living children of IUD acceptors in Tunisia and Kenya is as high as that of vasectomy acceptors in other countries."

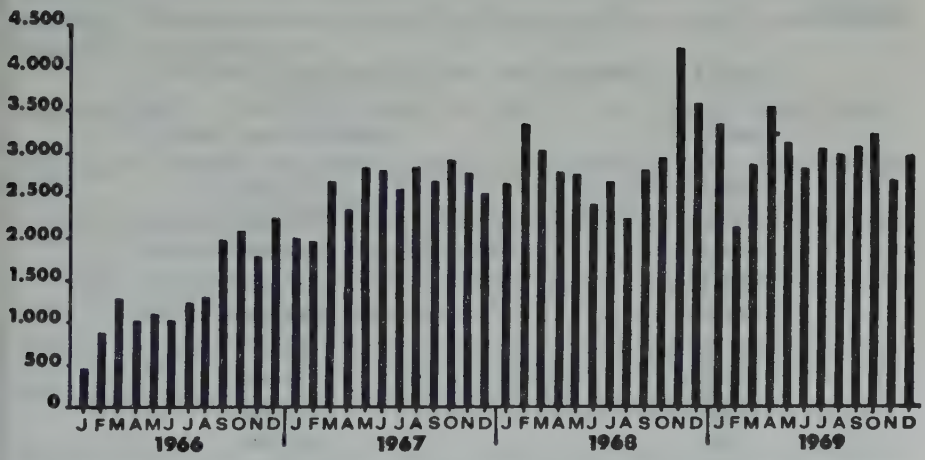
V. EVALUATION OF PROGRAM PROGRESS THROUGH STUDY OF TRENDS IN ACCEPTORS

Most programs, with an eye to evaluation, compile monthly data on the number of acceptors by method accepted. Data on the characteristics of acceptors such as age or number of living children are usually compiled less frequently. The main difficulty in using data on acceptors and their characteristics to study the progress of the program consists, as already indicated, in securing the appropriate base data against which those referring to acceptors can be related. This is particularly important in programs where acceptance rates



Source: *Taiwan's Family Planning in Charts*, Taiwan Population Studies Center, March 1967.

Fig. 2.1. *Loop acceptors by month*



Source: Singapore Family Planning and Population Board, *Fourth Annual Report*, 1969.

Fig. 2.2. *New family planning acceptors, 1966-1969*

are high. In the Singapore program, for instance, the annual acceptance rate stood at around 14 per cent of all married women in the age-group 20-44 years; in such circumstances the number of eligible women from whom new acceptors can be drawn diminishes rapidly within a few years of program operation, despite the continuous addition to the "eligible" group through new entrants into the age-group. Since acceptance varies markedly with characteristics such as age or number of living children, the study of trends in acceptance rates with respect to such characteristics is desirable but presents even more difficulties in securing data than the study of the total number of acceptors. For routine analysis, changes in the base population or its characteristics are usually ignored and the acceptor data for specific periods—monthly, quarterly or yearly—are assumed to be comparable.

TRENDS IN NUMBER OF ACCEPTORS AND SELECTED CHARACTERISTICS

NUMBER OF ACCEPTORS. The trend in monthly acceptors is represented by a line diagram as is usually done in Taiwan (Figure 2.1) or by a bar diagram as in Singapore (Figure 2.2).

Representation of the data through tables becomes more expressive when the data are shown by quarter or by full year. The annual numbers of new acceptors in Hong Kong (Table 2.8) for the period 1961 to 1970 show a rapid increase between 1963 and 1964, reaching a peak load in 1965. This was followed by a decline and then an increase.

Table 2.8. *Number of new cases in Hong Kong, 1961-1970*

Year (a)	*Number of new cases*
1961	12,299
1962	13,046
1963	14,861
1964	21,920
1965	35,946
1966	23,031
1967	19,292
1968	26,588
1969	30,886
1970	30,470

(a) April 1-March 31.

Source: Twentieth Annual Report, 1970-1971. The Family Planning Association of Hong Kong.

The fluctuations have been partly explained by Freedman *et al.* (1969). The trends in the number of acceptors for a few countries during the period 1967 to 1971 are given in Table 2.9.

Table 2.9. *Number of acceptors during 1967 to 1971 for selected countries (expressed in thousands)*

Year	Ecuador	Egypt	India (b)	Iran	Malaysia (West)	Singapore	Turkey
1967	3.3	141	2,985	10.4	20.7	31.7	47
1968	2.9	134	3,115	122.4	74.9	36.4	68
1969	2.6	148	3,411	230.5	70.6	37.0	77
1970	5.2	206	3,903	299.0	56.0	26.6	66
1971	12.3 (a)	221	3,799	385.0	55.1	21.7	51

(a) Data for 1971 are 95 per cent complete.

(b) Department of Family Planning, *Monthly Statements on the Progress of Family Planning Programme In India*, May 1971, p. 5, Ministry of Health and Family Planning, New Delhi.

Source: Dorothy Nortman, "Population and family planning programs: a factbook". *Reports on Population/Family Planning*, No. 2, September 1972, pp. 57-61.

In Singapore, the number of new cases increased gradually from 1787 in 1952 to 9845 in 1965, during which period family planning work was under the Singapore Family Planning Association. From January 1966, when the national program started, the number of new family planning cases showed a phenomenal increase, the total for 1966 being 30,410 (see Figure 2.2). As can be seen from Table 2.9, the number continued to increase until 1969 and declined subsequently. The countries covered in the table show a variety of trends in levels of acceptance.

ACCEPTANCE BY AGE AND NUMBER OF LIVING CHILDREN.

Ross *et al.* (1972) have drawn the following tentative conclusions:

(i) As program operation continues, the acceptance rates among younger women appear to rise.
(ii) Higher proportions of women with few children enroll as program operation continues.

Further data must be accumulated and analyzed before these plausible statements can be accepted as widely applicable.

ACCEPTANCE BY METHODS. Some programs place emphasis on one method only. In Iran and Western Malaysia, as stated earlier, the programs are directed mainly to the distribution of oral pills.

In such circumstances a study of the trend of acceptors through all program methods is tantamount to a study of the trend of acceptors of the main method offered.

In programs where several methods are offered, interest centers not only on the intake of the different contraceptives with time but also on the relative proportion of acceptors who chose different methods. In Tables 2.4 and 2.5 the percentages of acceptors in different years in Singapore and Hong Kong were presented. The Singapore program began stressing oral pills at the expense of the IUD around 1966 and this is well reflected in the sudden increase of pill acceptors in 1967. In 1969 there was evidence that acceptance of condoms, which is the next most popular method to the pill, was on the increase. The increasing popularity of the oral pill in Hong Kong at the expense of the IUD and the condom is brought out by the figures given in Table 2.5. The publication by Dorothy Nortman (1972) gives basic data which permit not only a study of the trend in acceptors of IUDs, oral contraceptives, sterilization and other program methods separately, but also the analysis of changes in the relative acceptance of different methods in a number of countries. It should, of course, be recognized that the trend in the acceptance of different methods would be affected by the relative emphasis placed by the program on the various methods offered, as well as the preference of the acceptor for these methods.

ANALYTICAL PROCEDURES

The ECAFE Regional Seminar (United Nations, 1970) felt that acceptance data by method might be compiled for each month and the trends investigated through "time series" analysis. The topics suggested for analysis were:

(1) total number of acceptors,
(2) acceptors in relation to total population,
(3) acceptors in relation to women of reproductive age,
(4) if appropriate data were available, acceptors in relation to number of eligible couples.

For routine appraisal of trends in acceptance, the calculation of relatively simple indices was recommended. These were percentage increase/decrease during the month as compared with:

(1) the preceding month,
(2) the monthly average of the preceding year,
(3) the monthly average over the preceding three years,
(4) the corresponding month of the preceding year.

It was suggested that similar indices might also be computed for quarterly acceptance figures.

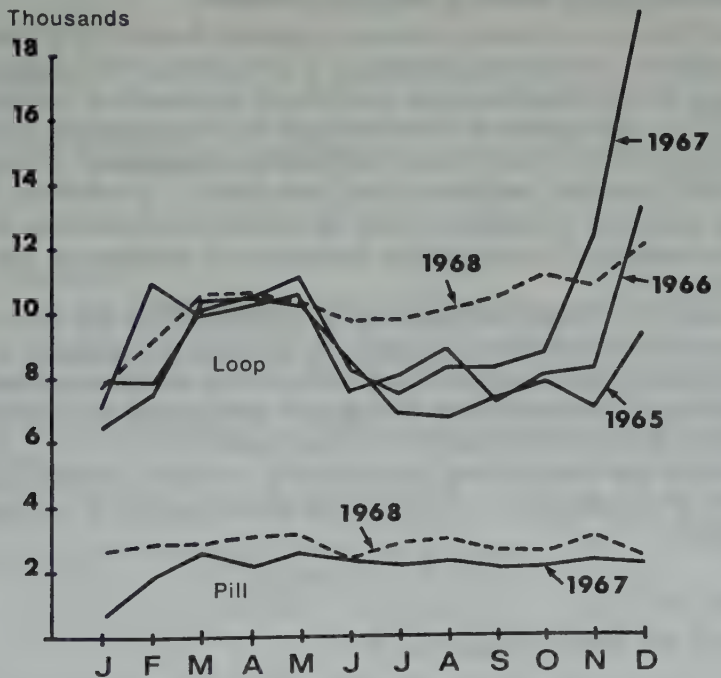
The seminar also recommended that for studying the trends in the characteristics of acceptors, acceptance data from year to year could be analyzed with respect to:

(1) age of wife,
(2) number of living children,
(3) education of wife and husband,
(4) rural-urban residential status of acceptor.

Where feasible it was suggested that trends in respect of religion, race and open birth-interval might also be added to the analysis.

In order to assess the changes in acceptance over time, the percentage distribution of total acceptors according to the methods accepted for the years for which data were available was recommended.

In general, only relatively simple analyses of time trends have been carried out so far in the evaluation of the progress of family planning programs. Even such analyses, however, have provided use-



Source: *Taiwan's Family Planning in Charts*, Taiwan Provincial Department of Health, 1969.

Fig. 2.3. *Monthly loop and pill acceptances*

ful guidance to program administration. To cite one example, in the early years of the Taiwan program a marked seasonal variation in the monthly loop acceptances was noticeable. In order to reduce this variation, the program provided added incentives to prospective acceptors at about the time when loop acceptances began to fall; such measures appear to have had a salutary effect as shown by the trend in monthly loop acceptances during 1968 (Figure 2.3).

MATHEMATICAL MODELS

Mathematical models can be used to test plausible hypotheses concerning the processes which underlie the trend in the number of acceptors. Chandrasekaran (1969) developed the following model assuming a two-stage process by which a person becomes an acceptor:

(i) A certain number of women are "prospective acceptors" at the start of program operation and an additional number become so during each month of the operation of the program. This number can be assumed to be α at the start of the program and β, a constant, to be added during each month of program operation.

(ii) There is a time-lag between a woman becoming a "prospective acceptor" and actually accepting a method from the program, the time lag being determined by the law of consecutive independent trials. That the chance of a woman who is a "prospective acceptor" at the beginning of any month becoming an acceptor during that month is a constant, assumed to be p less than 1. It is further assumed, for purposes of simplicity, that no woman who becomes a "prospective acceptor" in any month becomes an acceptor in the same month.

These two stages are realistic in the sense that the educational and promotional effects of a program lead to the creation of "prospective acceptors", while contact with the field staff and accessibility to service points determine the speed with which a "prospective acceptor" secures a method from the program.

Under the assumptions made above, it can be shown that the number of acceptors y_t in the t^{th} month after the start of the program is given by

$$y_t = \beta - (\beta - \alpha p)\, q^{t-1} \qquad (1)$$

where $q = 1 - p$

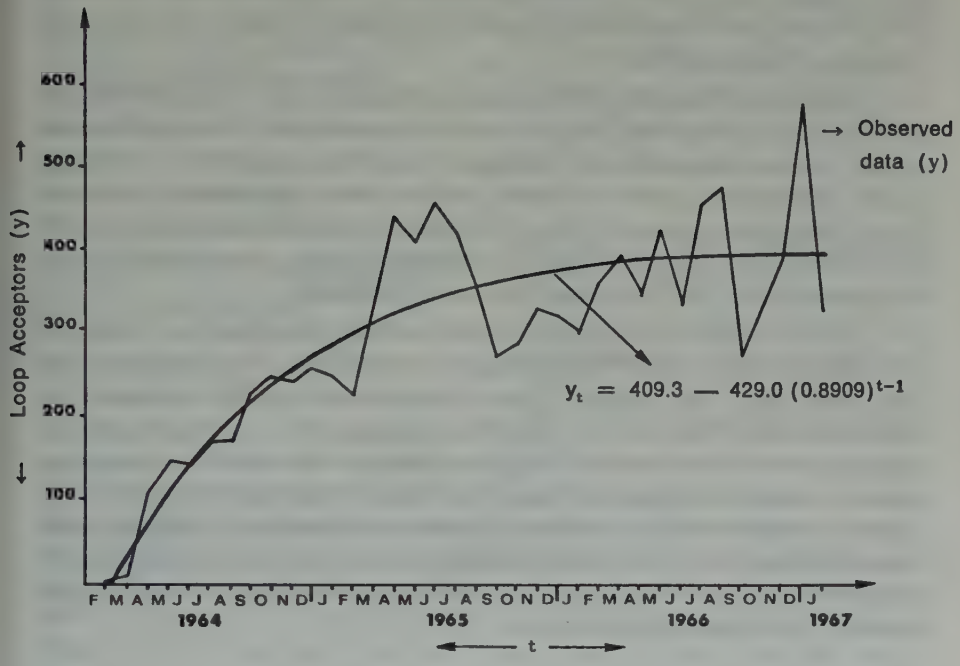
if $\beta > \alpha p$, the equation takes the form

$$y_t = K + a\, b^{t-1} \qquad (2)$$

where K is positive, a is negative and $b < 1$.

This model was fitted to the data on the monthly number of IUD acceptors in Kaohsiung county during the three-year period February 1964 to January 1967. (4) The data on monthly acceptors and the fitted curve

$$y_t = 409.3 - 429.0 \times (0.8909)^{t-1} \qquad (3)$$


Loop Acceptors (y)
600
500
400
300
200
100
→ Observed data (y)
y_t = 409.3 — 429.0 $(0.8909)^{t-1}$
F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J
1964
1965
1966
1967
t

Source: Data taken from Monthly Report of Taiwan Family Planning Program, Taiwan Population Studies Center, Taichung.

Fig. 2.4. *Monthly loop acceptors in Kaohsiung county, Taiwan*

(4) The fitting was done with the assistance of Dr. A. Mathai of the Mekong Office, ECAFE, Bangkok. As a first step b was estimated using the three sectoral totals $\Sigma_1 y$, $\Sigma_2 y$ and $\Sigma_3 y$ (where $\Sigma_1 y$ is the number of acceptors for the period February 1964 to January 1965 and similar definitions apply to $\Sigma_2 y$ and $\Sigma_3 y$) and the formula

$$b^{12} = \frac{\Sigma_3 y - \Sigma_2 y}{\Sigma_2 y - \Sigma_1 y}$$

The values of a and K were estimated by the formulae

$$a = (\Sigma_2 y - \Sigma_1 y) \frac{b-1}{(b^{12}-1)^2}$$

$$K = \frac{1}{12}\left[\Sigma_1 y - \frac{(b^{12}-1)}{(b-1)} a\right]$$

Other methods of fitting are being explored.

On the basis of the fitted curve which appears to reflect the trend in the monthly acceptors pretty well, it could be stated that over the three-year period, the number of women who became "prospective acceptors" each month was 409. Also, the chance of a woman who was a prospective acceptor at the beginning of a month becoming an acceptor during that month was 0.1091. Such estimates can provide a program administrator with summary indices of important aspects of the program.

The above discussion should be construed only as an illustration of new approaches that can be used to understand the significance of the trends observed in the number of monthly acceptors. Although in the illustration given this trend, *prima facie*, indicates an increasingly better performance in the first year of program operation, the fit as given by the model indicates relatively static conditions of performance during the three-year period considered.

The two-step model outlined above has considerable flexibility and can be used under more stringent conditions. For instance, the condition that β is constant over a time period can be replaced by the assumption that the "prospective acceptors" in a month will be dependent on the number of eligible women at the end of the previous month. Initial attempts at fitting the model described above to the data on monthly IUD acceptors from other counties in Taiwan have revealed the need for using different values of the parameters in different periods. Such changes in parameters are of great significance to program administrators.

VI. USE OF ACCEPTOR DATA FOR IMPROVEMENT OF ADMINISTRATION AND MANAGEMENT

Assessment of program performance and progress using acceptor data as described in the last two sections is directed to the improvement of administration and management of family planning programs. The methods described in these sections depend largely on data available through routine records such as the individual client card or the coupon. More detailed assessment can be made by analyzing the data on acceptors obtained from such records with related information obtained from other sources. The illustrations which follow exemplify some of the approaches which have been made in this direction.

In a study of the effectiveness of clinic services in Greater Bombay, Chandrasekaran and Kuder (1961) devised two indices to assess the rate at which new cases were enrolled. These were the number of new cases enrolled per hour a clinic was open and the number of new cases enrolled per professional hour, as defined below.

$$\text{Number of new cases enrolled per hour a clinic was open} = \frac{\text{Number of new cases enrolled during a calendar year}}{\text{Number of hours clinic was scheduled to be open during a week} \times 52}$$

and

$$\text{Number of new cases enrolled per professional hour} = \frac{\text{Number of new cases enrolled during a calendar year}}{\text{Number of hours clinic was scheduled to be open during a week} \times \text{number of professional staff sanctioned for the clinic} \times 52}$$

These indices provided the basis for comparing the performance of a number of clinics which had different work schedules, different staff provision and worked under different auspices, such as municipal corporations, Family Planning Association of India, and other voluntary organizations. The analysis showed, for instance, that the highest case load per hour was reached in municipal clinics when the clinic was open for four hours a week. A similar finding was also recorded by clinics run by "other voluntary organizations". In these two types of clinic there was not only a larger number of new cases but also more rapid attendance per hour.

Studies undertaken in connection with the Taiwan Family Planning Program demonstrate the variety of ways in which data on acceptors can be analyzed in combination with other types of data in order to provide useful leads to the program administrator in improving its administration and management. The additional data used have been obtained from specially planned investigations, from the program itself, or from other routine sources. Two of the specially planned investigations which will be mentioned here relate to the assessment of the postpartum mailing program and of the "free offer" for a limited time, in increasing acceptance. The attraction of the postpartum mailing program lies in its potential effectiveness in increasing acceptance at a much lower cost than through a program

of routine visits by fieldworkers. The effect of the "free offer" for a limited period of three months has been a large increase in the number of acceptances—an increase of "100 to 175 per cent as compared with the preceding three-month period". Careful analysis of the trend in acceptances for a period covering six months after the end of the free offer resulted in the conclusion that the increase in acceptance during the "free period" was not "a borrowing of cases from the future" (Freedman and Takeshita, 1969).

In another study the acceptance rate was related to program input using geographical area as the unit for comparison. The family planning worker ratio defined as "months of work in 1964 and 1965 by postpartum health workers and village health education nurses per 100 married women aged 20-44" was the index used for program input. Nine other factors likely to affect the acceptance rate—doctor/population ratio, percentage of females with at least primary education, total fertility rate, change in total fertility rate, percentage of the male labor force engaged in agriculture and fisheries, net migration rate, logarithm of population density, distance to nearest city of 50,000 inhabitants or more and crude death rate—were also taken into account in studying the relationship of family planning worker ratio to acceptance rate through regression analysis. As data for these nine factors were obtained from non-program sources, the study in fact used information from program as well as external sources. The analysis brought out a strong positive relationship between program input and acceptor rate (Freedman and Takeshita, 1969).

Acceptor data were used by Lapham and Mauldin (1972) to review and evaluate programmatic considerations on an international basis, taking into account the information available from 20 countries with national family planning programs. The output of the programs and statistical measurements of acceptors and users were utilized. Since the present discussion relates to acceptors, only the measurements relating to them will be mentioned here. These were:

(1) acceptors as a percentage of married women of reproductive age (15-44 years) defined earlier in this chapter as acceptance rate;
(2) ratio of acceptors to net one-year gain in number of married women of reproductive age (15-44 years);
(3) annual acceptor target and per cent of annual target achieved;
(4) cost per acceptor.

The programmatic criteria used were fifteen in number. These were

largely designed to obtain information on the presence or absence of certain traits including the following

(i) Does the government have a population policy that includes fertility reduction?
(ii) Do the customs and legal regulations facilitate importation of contraceptive supplies that are not manufactured locally?
(iii) Does the Ministry of Health carry out a vigorous effort to provide family planning services to all married women of child-bearing age in the country?
(iv) Are adequate family planning training facilities for health personnel available and utilized?

Various social, economic and health criteria were also used in the analysis.

The statistical measurements relating to acceptors can by themselves be illuminating. Taking for instance the cost per acceptor, Lapham and Mauldin (1972) state:

> "Bearing in mind the reservations and difficulties associated with cost data, the cost per acceptor varies from US $4.00 or less per acceptor in Indonesia and Singapore to over $20.00 per acceptor in Turkey, India and Tunisia. For some countries no cost data are available. The 1969 Tunisia $39 per acceptor and the $0.15 expended per capita indicate a well-financed program."

They add,

> "The high cost per acceptor for some countries raises a major question. It is probably reasonable to assume that neither individual governments nor donor agencies are likely to continue to support levels of $20 to $40 per family planning acceptor indefinitely, even though these costs may be reasonable in terms of future savings in expenditure for education, health services, creation of new jobs and so on, resulting from averted births."

Lapham and Mauldin studied the relationship between acceptor rates and the ratings on the fifteen programmatic criteria mentioned earlier. For this purpose they divided the acceptor rates into three groups (5.0 per cent and over was classified as high; 3.0-4.9 per cent and under 3 per cent made up the two lower groups). Each programmatic criterion was classified into four groups—"yes", "qualified yes", "partially" and "no"—according to the descending order of the intensity of the procedural and supportive measures that pertained to the criterion. Leaving aside four programmatic criteria that were similar in the programs of all countries considered, the authors found that countries with higher acceptor rates had a larger percent-

age of "yes" answers when the ratings on the eleven other criteria were taken together. Taking the countries in the "high" acceptor group, 65 per cent of the ratings were "yes" and only 10 per cent were rated as "partially" or "no". The remaining 25 per cent were in the category "qualified yes". In the countries with medium acceptor rates (3.0 to 4.9 per cent), 22 per cent of the ratings of the programmatic criteria were "yes" and 39 per cent "partially" or "no". In the remaining group of seven countries with low acceptor rates, there were few "yes" or "qualified yes" ratings; 70 per cent of the ratings were "partially" or "no".

Two of the more striking differences were for the criteria "contraception readily and easily available" and "vigorous effort to provide family planning services to all married women of child-bearing age in the country". The higher acceptor rate countries tended to have contraception readily and easily available in both public and private sectors and they tended to carry out vigorous efforts to provide all couples with services; the lowest acceptor rate countries fell almost entirely into the "partially" category on these two criteria. The extensive use of fieldworkers and mass media also showed strong differences between the higher and lower acceptor-rate groups.

Lapham and Mauldin recognize the limitations of their findings, based as they are on a small number of countries. They conclude:

> "The general picture of these 11 criteria strongly suggests activities and approaches the countries in the lower acceptor-rate groups might well consider when their goals and intentions are to increase the use of contraception and planning of family size."

VII. USE OF ACCEPTOR DATA FOR ASSESSING THE IMPACT OF PROGRAMS

Acceptance of services and/or advice from a family planning program is an important step in the utilization of the facilities provided by the program for achieving individual goals in respect of family formation as well as national goals set by the program. Whether the achievement of individual goals will ensure the achieving of national goals and vice versa is debatable but most national programs, as pointed out in Chapters 1 and 3, place greater emphasis on national goals.

These goals are in large part directed towards effecting reductions in fertility. As pointed out earlier, changes in fertility will result not only because of program efforts but also because of changes in dispositions towards family size arising from social and

economic trends occurring in society. However, in assessing the impact of a program, the first impulse and certainly a more direct endeavor is to depend as much as possible on the information which it readily provides. Thus acceptor data become important in the evaluation of the effects of a program.

In the earlier sections of the present chapter, we have discussed the use of acceptor data for evaluating program progress. Reference has also been made to efforts that are made to compare the number of acceptors with the "targets" set for acceptors as a method of relating achievement to goals. The setting of targets for acceptors has come into practice in the past few years, and Chapter 3 describes the targets for several countries and discusses some of the problems that arise in relating the targets to anticipated reductions in birth rate.

Acceptance of a method signifies also the beginning of use of the method in cases such as sterilization and IUD; in most other cases acceptance merely implies the possibility of future use by the acceptor. The length of use of a method depends on the purpose for which it is used—whether it is intended for spacing of births or for their limitation. It also depends on the extent to which the couple is satisfied by the use of the method as judged by the absence of physical discomfort, side effects, etc. New procedures for assessing the length of use which have been developed in recent years utilizing the records of acceptors as the starting point are discussed in Chapter 5.

Use of the method provided by the program is an important intermediate stage between acceptance and the postponing or prevention of births. Methods evolved for estimating births averted by the program utilizing the data on acceptors and the length to which the method accepted by them has been used are discussed in Chapter 6. Since such discussion involves a quantitative appreciation of some of the biological elements involved in human reproduction, a background to this topic is provided in Chapter 4.

Acceptor data have been used to investigate whether the fertility changes observed after the initiation of a program can be attributed to the influence of trends in social and economic variables which might already have been affecting fertility and what additional effect the program might have produced. Two methods of investigating the program and non-program effects are dealt with in Chapters 7 and 8.

In summary, acceptor data are extremely useful in assessing not only the performance and progress of family planning programs but also their impact on fertility.

REFERENCES

Chandrasekaran, C.

1969 "Family limitation in South and South-East Asia", International Population Conference, Vol. II, International Union for the Scientific Study of Population, London.

Chandrasekaran, C. and Malini Karkal.

1972 "Continuation rate, use-effectiveness and their assessment for the diaphragm and jelly method", Population Studies, Vol. XXVI, No. 3 (November), London.

Chandrasekaran, C. and Katherine Kuder.

1961 "The organisation and case-load of family planning clinics in Greater Bombay", Report of the Fourth All-India Conference on Family Planning, Family Planning Association of India, Hyderabad.

Chow, L. P.

1968 "Evaluation of the Family Planning Program in Taiwan, Republic of China", Journal of the Formosan Medical Association, Volume 67, No. 7 (July).

Freedman, Ronald, D. N. Namboothri, A. Adlakha and K. C. Chen.

1969 "Hong Kong: The continuing fertility decline, 1967", Studies in Family Planning, No. 44 (August), The Population Council, New York.

Freedman, Ronald and J. Y. Takeshita.

1969 Family Planning in Taiwan: An Experiment in Social Change, Princeton, Princeton University Press.

Hong Kong Family Planning Association.

1971 Twentieth Annual Report, 1970-71.

Lapham, Robert J. and W. Parker Mauldin.

1972 An Assessment of Family Planning Programmes, OECD Development Centre. Also abstracted in Studies in Family Planning, Vol. 3, No. 3 (March), New York, The Population Council.

Nortman, Dorothy.

1972 Population and Family Planning Programs: A Factbook, No. 2 (Fourth edition), New York, The Population Council.

Planned Parenthood Federation of America.

1972 Family Planning Perspectives, Vol. 4, No. 3 (July), New York.

Ross, John A., Adrienne Germain, Jacqueline E. Forrest and Jeroen Van Ginneken.

1972 "Findings from Family Planning Research", Reports on Population/Family Planning, No. 12 (October), New York, The Population Council.

Ross, John A., Frederick F. Stephan and Walter B. Watson.

1968 A Handbook for Service Statistics in Family Planning Programs, New York, The Population Council.

Singapore Family Planning and Population Board.

1969 Fourth Annual Report.

Takeshita, John Y. and Ronald Freedman.

1967 "Measuring acceptances in a family planning program; the decomposition of rates by eligibility criteria", Demography, 4.

Taiwan Population Studies Center.

1964-67 Monthly Report of Taiwan's Family Planning Program.

1967 Taiwan's Family Planning in Charts (March).

Taiwan Provincial Department of Health.

1969 Taiwan's Family Planning in Charts (June).

Tietze, Christopher and Sarah Lewit.

1973 "Recommended Procedures for the Evaluation of Intrauterine Contraception", Studies in Family Planning, Vol. 4, No. 2, New York, The Population Council.

United Nations.

1969 Assessment of Acceptance and Effectiveness of Family Planning Methods, Report of an Expert Group, Asian Population Studies Series, No. 4, E/CN. 11/882.

1970 Evaluation of Family Planning Programmes, Report of a Regional Seminar, Asian Population Studies Series, No. 5, E/CN. 11/936.

1973 A Comparative Study of Service Statistics Systems for Evaluation of Family Planning Programmes in the ECAFE Region, ECAFE, Bangkok (mimeographed).

Wishik, Samuel M. and Kwa Hwa Chen.

1973 Couple Year of Protection: A Measure of Family Planning Program Output, Manual of Family Planning and Population Programs No. 7, International Institute for the Study of Human Reproduction, Columbia University, New York.

Wolfers, David, ed.

1970 Post-Partum Intra-Uterine Conception in Singapore. Excerpta Medica, Amsterdam.

Chapter 3

ACCEPTOR TARGETS *

by

John A. Ross

I. INTRODUCTION

Of the developing countries that have decided to reduce their population growth rates, about thirteen have worked through a set of calculations as to how many family planning acceptors are required. These calculations have never been collected and appraised, and it seemed useful to put such a review alongside the more analytic models presented in this volume for assessing births averted by a program.

Those models offer conceptual guidance to particular aspects of target setting in a more scientific way, and with better data, than are typically available to program planners and administrators. They point the direction in which the field might profitably move. In turn, the descriptions of current practice below may outline actual situations which future scientific work may wish to consider.

The plans were of course drawn up in situations of greatly differing data resources. Moreover, they had to be tailored to the circumstances of each country—its particular administrative framework, its readiness to be explicit about targets, its service delivery system. Conceptual differences also operate, affecting the meaning of acceptance, continuation, births prevented and other elements of the plans. Certainly no single plan can ever serve for all countries; nevertheless estimates must inevitably be made of the numbers of

* Portions of this chapter appeared earlier in *Seminar on Demographic Research in Relation to Population Growth Targets*, 3-9 April, 1973, University of the West Indies, St. Augustine (Trinidad and Tobago), Committee for International Coordination of National Research in Demography (CICRED), Series S., No. 1.

Table 3.1. *Proportions female of childbearing age and percentage married: Selected countries*

Country	Year	Females per 1,000 population			Married females per 1,000 population			Percentage of females married	
		15-49	15-44	20-44	15-49	15-44	20-44	15-44	20-44
Algeria	1966	206	190	145	(a)	143	124	75	86
Ceylon	1963	222	203	156	144	129	122	64	78
Chile	1960	242	219	169	134	117	113	53	67
Colombia	1964	227	209	156	127	115	107	55	69
Egypt	1960	217	195	164	164	147	138	75	84
India	1961	228	209	170	193	180	152	86	89
Indonesia	1964-65	241	221	183	(a)	167	(a)	76	(a)
Iran	1966-68	200	186	145	168	156	137	84	94
Iraq	1965	199	182	143	(a)	128	(a)	70	(a)
Kenya	1962	234	219	172	165	153	130	70	76
Malaysia (West)	1957	222	203	155	161	147	130	72	84
Mexico	1960	240	222	159	141	128	117	58	74
Morocco	1960	224	210	180	180	171	156	81	87
Nepal	1961	249	229	187	211	197	167	86	89
Pakistan	1961	209	193	153	171	160	132	83	86
Peru	1961	227	211	159	137	123	116	58	73
Philippines	1960	232	213	160	141	126	119	59	74
South Korea	1966	226	207	162	(a)	130	(a)	63	(a)
Thailand	1960	232	213	166	146	132	126	62	76
Turkey	1965	218	204	161	(a)	158	(a)	77	(a)
Venezuela	1961	222	203	155	129	117	107	58	69
Zaire	1955-58	(a)	251	214	(a)	206	189	82	88

(a) Unavailable

acceptors needed for a given demographic goal, and it is instructive to compare the ways so far used to do so.

Before turning to the plans themselves, a brief picture of the fundamental situation they all face may be useful. An overall look at the population to be served is relatively easy. Table 3.1 shows that in a developing country population of 1,000 there will be:

200-250 females aged 15-49, 130-200 of them married
180-230 females aged 15-44, 115-185 of them married
140-190 females aged 20-44, 110-160 of them married.

Adjustments can be made to these numbers: some couples will be naturally sterile, the exact percentage being highly dependent on age and ranging from a low of around 3 per cent at female ages 20-24 to over 30 per cent for ages 40-44; Henry's work (1953, 1961a, b) has been used to obtain an approximation. If surveys have gathered the appropriate data, a further subtraction can be made for those who have had sterilization operations, and an addition can be made for those in cohabitation as distinct from formal marriage. The end result of this will be the determination that 10-15 per cent of the population is eligible, with their husbands, for program contact. (Those currently pregnant are included, as in any long range plan they are eminently eligible).

Supposing that the crude birth rate is 40 per thousand, it is these 100 to 150 females who are responsible for the 40 births each year. The program might start by saying that the crude birth rate can be reduced by one-fourth, i.e., from 40 to 30 per thousand, by recruiting about one fourth of the child-bearing population just identified. However, besides being a very large undertaking, this would by no means guarantee the desired result.

First is the difficult problem of acceptor selectivity. The correlates of greater effect are tied to the *program's uniqueness* (did the acceptors use birth control before? are they living in an area where alternative services are available?), the *potential fertility* of acceptors (are they young? of low parity? with a recent birth?) and the *methods offered* (how many? how good is continuation?). There are innumerable complexities: sterilization offers the best continuation but is not attractive to younger couples; a program may start more readily where it is not unique—where the public is already accustomed to private birth control and commercial methods.

Moreover, the program must work under constraints related to the family cycle. About 20 per cent of all births will be first births and actively sought by the couple. A somewhat lower percentage will be

second births ([1]). In general, acceptance comes from the broad middle group—not the lowest parities nor the youngest couples, nor the oldest and the sterile, but those in between. There is usually a substantial spread as well—appreciable numbers do accept after the first two births, and surprising numbers come from the upper age and parity categories. Therefore the focus must be a broad one, both because of this spread and because different sub-groups move at different rates towards new attitudes and practices.

Fortunately, the acceptors to a certain extent select themselves in such a way as to bring into the program those with a somewhat greater proclivity to future births. Clearly couples who know themselves to be sterile will not accept, nor will those who are already using highly effective methods, and this immediately creates a differential. Especially in the early period, couples taking up program contraception tend to be those to whom other sources are unknown or unattractive, and therefore unlikely to have been used.

Additional complexities arise from projections into the future. In young programs these are often omitted as it will take up to five years just to reach those initially identified. But the time flow has to be faced, and most of the plans described in this chapter have done so. Large numbers of new entries each year come up through the pyramid of marriage and early births, large numbers leave through aging and broken marriage and, in some countries, a very large young cohort arrives to worsen the age structure of the eligible group. Also, many acceptors move from one method to another, confusing the count of how many couples have been covered by the program. All of this introduces new and difficult calculations for the "maintenance" aspect of the program.

Thus the estimation of needed acceptors for a given fertility reduction can challenge the most ingenious planners. The complexities involved have been handled very differently in the different plans, and we now present the individual descriptions. These cover fourteen plans in thirteen countries, arranged in alphabetical order.

([1]) First births as a percentage of all births run illustratively as follows:
West Pakistan, 1962: 16
East Pakistan, 1962: 17
Thailand, 1960-64, 1966: 21-22
The percentages for second births are 17, 15 and 18, respectively. See Ross (1970) and Thailand, *Statistical Yearbook* and *Public Health Statistics.*

II. METHODS USED IN ESTIMATING TARGETS COUNTRY EXAMPLES

CEYLON

Ceylon's goal has been to reduce its crude birth rate from 36 in 1966 to 25 in 1975, which would lower the population growth rate to 1.7 per cent and increase the period in which the population would double from about 28 to 41 years. A set of calculations (Perera *et al.*, 1968; Wright, 1974) done in 1968 estimated that 115,000 new acceptors would be needed annually to meet this target, or 690,000 from January 1969 through 1974.

The reasoning went as follows:

(a) Marital fertility was about 250 per thousand women (1963 data), suggesting that it would take four married couples practicing 100 per cent effective contraception for a year to prevent one birth. Assuming an unchanged mortality pattern, keeping the same proportion of females married at each age, using 1965 age-specific birth rates, and projecting the number of females to 1975, about 150,000 births would have to be prevented in 1975 by birth control both inside and outside the program. At the 1:4 ratio above, the prevention of 150,000 births in 1975 requires at least 600,000 effective and continuing users.

(b) However, users are a residual sub-group within acceptors. Studies elsewhere on the IUD show only a 70 per cent fall in fertility among acceptors in the two years after acceptance, and this comes partly from the use of substitute methods, including abortion, after IUD termination. Therefore the 600,000 must be raised to 860,000 (600,000/.7), and even this requires that all other methods are as effective as the loop and that experience beyond the two years matches the 70 per cent figure.

(c) In practice, continuation rates for pills and condoms will be inferior, and some acceptors will be merely spacing. As a reasonable guess 70 per cent should be reduced to 60 per cent, producing a figure of about 1,000,000 users to prevent the original 150,000 births annually by 1975. Assuming that the percentage of women aged 15-49 who are married is the same in 1975 as in 1963, the 1,000,000 will represent 45 per cent of all married couples.

(d) The private sector will meet perhaps 25 per cent of this goal; hence, the 1,000,000 falls to 750,000.

(e) The 750,000 must be supplemented to offset users who, during the period 1968-75, arrive at menopause or become divorced or widowed. If these are 1 per cent of the 750,000 annually repeated over 7 years, the total should be raised by 50,000 to 800,000.

(f) Subtracting the 100,000 acceptors recruited during 1966, 1967, and 1968, (prior to the 1969-74 period under calculation) leaves 700,000 or 115,000 per year. This is an annual target of about 5.3 per cent of married couples. Actually 1969 acceptors numbered only about 60,000, a shortfall of one half, and performance over the years through 1972 ran at about 2-3 per cent of MWRA (married women of reproductive age) (4.9 per cent in 1973).

The Ceylon plan contains features of special interest, and tries to handle several important factors. Further comments are made in the "Appraisal Section".

GHANA

Ghana aims to reduce the crude birth rate by 1 point per year, to keep pace with an expected decline of the same amount in the crude death rate. The straightforward six-column table below (Gaisie and Jones, 1970) is produced by estimating that

(a) about one-third of acceptors in a given year will discontinue use that same year,
(b) about one-third of the previous users who enter a year will discontinue in that year, and
(c) in each year, one birth is prevented for every four users of effective methods active at the end of the year.

Taking Ghana's population at 10 million, the 1974 crude birth rate would be reduced by 5 points, i.e., from 47 to 42.

Table 3.2. *Ghana: Planning targets for increasing the use of contraceptives (Figures are in thousands)*

Fiscal year	*Users at start of year*	*Acceptors during year*	*Total number of users during year*	*Users discontinuing during year*	*Users at end of year*	*Estimated births prevented*
1970		54	54	18	36	9
1971	36	78	114	38	76	19
1972	76	98	174	58	116	29
1973	116	118	234	78	156	39
1974	156	144	300	100	200	50

INDIA

The available documents on the Indian program (S.C. Jain, 1971) permit a description of calculations on the likely demographic effect

of certain acceptor targets, rather than the manner of estimating acceptor targets to meet given demographic goals.

As of early 1971 the goal was to reduce annual population growth from 2.5 per cent to about half that level, implying a crude birth rate (CBR) of about 22-23. Targets by method of birth control (sterilization, loop, conventionals) were stated not in absolute numbers but on a per thousand population basis, necessitating various adjustments, not described here, in the translation to effects on total numbers of births and population size.

The calculations were performed within the context of a component population projection, taking account of life-table mortality and changing age structure. For sterilization and loop acceptors, age distributions were obtained from various local studies and were assigned the marital age-specific fertility rate (MASFR) schedule prevailing for the general population, derived from data taken from the 14th round of the National Sample Survey. Among acceptors, subtractions were made for mortality and, in the case of the IUD, for termination of use (with retention after years 1, 2, 5 and 10 of .656, .488, .200 and 0 respectively).

Acceptors of conventional contraceptives are given a potential fertility rate of 250 births per 1000 couples per year, equal to that estimated for all couples of reproductive age ([2]). This assigns to conventional acceptors the age distribution of the general population. Use-effectiveness for conventionals is taken at 60 per cent, the usual figure for condoms, since these make up the majority of conventionals. Thus a regular user of conventionals prevents .15 births ($.60 \times .25$) per year of use. It is not clear how "regular users" are distinguished within the acceptor counts in the program service statistics.

Finally, a conversion is necessary to translate the births prevented as calculated above from an acceptor basis to a fiscal year basis. In connection with this, certain adjustments are mentioned (though not explained) pertaining to (a) non-pregnancy at acceptance, (b) the assumption "one year of exposure to the contingency of contraception" in deriving "births saved in successive years of use" and (c) the nine-month lag from conception to birth.

The plan projects a remarkable program effect upon the crude birth rate, showing a fall of over 14 points, i.e., from 36.5 to 22.1 in the ten years ending in fiscal year 1978-79. The annual acceptor targets

([2]) If the crude birth rate is 42, and there are 180 couples per 1000 population as in the 1961 census, then the couple fertility rate, 42/180, equals .23, close to the .26 obtained by an independent method; hence .25 is adopted.

are ambitious, rising from 17 to 35 per 1000 population per year. Moreover, the proportion of sterilizations is high at one third to one quarter of all acceptors, IUD continuation rates are high, and virtually no discounts are made for useless IUD wearing time. (On the other hand, all secular influences such as rising age at marriage and private birth control are neglected.)

INDONESIA

In the basic Indonesian plan (United Nations, 1970), acceptances were targeted at 6 million over five years (1971-75), increasing annually in the ratio 1:2:4:6:7, with IUDs, pills and conventionals going to 50, 30 and 20 per cent of new acceptors respectively. Each year's acceptor cohort for each method was moved forward under a decay curve, and simple totals were drawn to obtain total users at the start of each year. A related calculation was performed to obtain a summation of woman-months of use. Under an assumed potential fertility rate of 333 (birth interval of 3 years) the number of woman-months was simply divided by 3 to obtain births prevented nine months later. The numbers rise over time, so that by the 1975-76 fiscal year, they would with a nine-month time lag reduce the crude birth rate of Java, Madura and Bali by 8.5 per 1000 or that for the whole of Indonesia by 5.6 per 1000.

The plan recognizes that what terminators do is critical: if they somehow continue protection the birth rate will fall more; if they do not it will fall less due to their concentration in the fecund part of the birth interval as they terminate use. Indeed, if unprotected they will then have a surge of births, rendering many of the births prevented as only births temporarily averted (postponed) (Chandrasekaran *et al.*, 1971). Unique among those reviewed, the Indonesian plan contains guidelines indicating the percentage of terminators of each method who should take up each substitute method. The plan attributes importance to the long lactation period of Indonesian women, implying that otherwise the potential fertility of .333 should be put higher. The presence of overlap between use and lactation, however, receives no attention. Helpful indirect effects of the program are mentioned, such as the motivation of the general public through the mass communication program.

JAMAICA

The International Bank for Reconstruction and Development (IBRD) Mission of 1972 (IBRD, 1972) to the Government of Jamaica developed the calculations shown in Table 3.3. Actual acceptances

Table 3.3. *Jamaica: Effect of the number of new acceptors on the birth rate*

Year	Number of new acceptors during year [a] (1)	Number of new acceptors in the program at the end of year [b] (2)	Number of acceptors in the program at the end of subsequent years [c] (3)	(4)	(5)	Total number of acceptors in program at end of year [sum of columns 2-5 [d]] (6)	Women-years of contraceptive use [e] (7)	Number of prevented births (8)	Projected population (9)	Cumulative reduction in the birth rate [g] (10)	Level of birth rate (1968= 34.2 per 1,000) [h] (11)
1969	28,870	20,209	5,641			25,850	19,310		1,863,700		
1970	19,014	13,310	11,115	3,554		27,979	26,915	4,828-6,437 [f]	1,890,710	2.6-3.4	31.6-30.8
1971	21,472	15,030	7,321	7,002	2,097	31,450	29,715	6,729-8,972	1,911,420	3.5-4.7	30.7-29.5
1972	20,928	14,650	8,266	4,612	4,131	32,812	32,131	7,429-9,905	1,935,280	3.8-5.1	30.4-28.7

(*a*) Assumed to be evenly spread over the whole year.

(*b*) These are the acceptors who joined the program during the course of a year and that remain at the end of the year with a 50 per cent continuation rate. Column (2) is approximately equal to 70 per cent of column (1).

(*c*) Acceptors remaining from previous years.

(*d*) These are the acceptors who remain in the program but who joined in years other than the current year.

(*e*) Average of the two preceding figures in column (6).

(*f*) Range given by a foregone fertility of acceptors of 3 to 4 prevented births per woman-year of contraceptive use. Col. (8) = Col. (7) divided by 4 and by 3, respectively, and lagged one year.

(*g*) Col.(10) = Col. (8) divided by Col. (9)×1,000.

(*h*) Col. (11) = 34.2 minus Col. (10).

Source: Jamaica National Family Planning Program. Report of First External Review. Report submitted to the Government of Jamaica, Washington, D.C., July 1972. Annex 6.

for each year were projected forward under decay curves and summed to obtain users as of December 31. Woman-years of protection within each year were made equal to the simple mean of users at the start and end of the year, a procedure that serves quite well when annual acceptances are reasonably level as they were in Jamaica.

Births prevented were calculated on a ratio to woman-years of use, ranging from 1:3 to 1:4, and were lagged by one year. These were then substracted from the births expected under a base-line crude birth rate of 34.2 per thousand and a projection of total population, yielding the new birth rates in column 11 of Table 3.3.

The procedures, while simple, ignore any trends in private birth control, marriage or age structure, as well as discounts on life table using time (to recognize mortality overlap with post-partum amenorrhea, and other factors discussed below) and the possibility of the post-termination surge of births. The conversion of births prevented to a rate (column 10) parallels that done for Indonesia, and permits a statement of program effect ignoring the other determinants. It suggests how much the program reduces the crude birth rate over the (unstated) level it otherwise would have been. Mortality is ignored.

REPUBLIC OF KOREA

1962-1971. The national program began in early 1962, offering vasectomy, condoms and other traditional methods including rhythm. In 1964 the loop became available for mass use, and a ten-year plan was written covering the period from January 1962 through December 1971 (Taek Il Kim, 1970). This plan charted the course of fertility, mortality and growth year by year through the ten-year period, first without a program and again with it. The plan included the following: (Taek Il Kim, 1972).

> "...that 100 married women would ordinarily have 25 births per year and that loops in place would suppress all 25. Originally, 20 per cent of loops were predicted to drop out in the first year and 10 per cent per year thereafter (including regular terminations plus deaths and menopause). Vasectomies were assumed to remain effective to December 1971, with deaths, menopause, and so on, disregarded. Conventional methods were discounted by 70 per cent, down to a 30 per cent level. All acceptance in each calendar year repressed no fertility until the next calendar year. Therefore, if 100 loops were inserted one year and 20 per cent dropped out, the fertility repressed in the next year was $100 \times 80\ \% \times .25 = 20$ births. For 100 vasectomies it was $100 \times .25 = 25$. For 100 couple-years of conventionals, it was $100 \times 30\ \% \times .25 = 7.5$. By projecting the targeted acceptance by year and by method, births prevented and couples protected were calculated easily. By December 1971, 31.6 per cent of couples would be protected through these government services, and it was assumed that another 13.4 per cent of married women aged 20-44 would find protection on their own, making 45 per cent in all.

"Using 1960 census data... population projections... and Japanese and Taiwanese experience, 1962 fertility crude birth rate was taken at 40 and estimated to decline naturally by .4 points per year (4 in 10 years to 36 in 1971), taking into account an estimated 13.4 per cent of couples practicing on their own. To this was added each year's program effect, translating the births-prevented figures into an additional decline in the crude birth rate, all with the end result of 27.0 birth rate by December 1971. Mortality was taken at 10.5 in 1962 and was expected to decline by .3 points per year (3 in 10 years, to 7.5 in 1971). The difference was a 1.95 per cent growth rate at the end of the plan period, a shade below 2.0 per cent.

"With only two or three revisions, this plan stood as the operating blueprint into 1970... The principal revision was to elevate the IUD target from 1 million to 1.8 million to recognize that IUD terminations would be higher than originally thought. Balancing this was the fact that no account was taken of induced abortions which by 1968 were ending at least 20 per cent of all pregnancies, or of pills, which were selling about 125,000 cycles per month commercially and going to about 240,000 couples in the program by December 1970. Also ignored was the rapidly rising age at marriage, which has removed many highly fertile young women from early child-bearing and lengthened the generational span."

1972-1976. Quite a number of draft plans, covering different time periods, have at one time or another been drawn up for the Korean program, and one of these (Watson and Smith, 1971) is of interest for this discussion even though it has no official standing. Technically it offers features not used in most of the other plans.

Its reasoning is that fertility is declining from the high and constant age-specific levels that would prevail without the presence of later marriage and birth control. The 1960 age-specific rates represent the high level from which the decline is occurring and the degree of decline follows the proportion of women protected by later marriage and by each birth control method. These fertility depressants had sufficient force as of 1970 (the time of writing the plan) to produce the age-specific levels then prevailing, and the absolute numbers of births being prevented are easily derived as the difference between total births under the 1960 and the 1970 rates applied to the 1970 age distribution. The various fertility depressants must be strengthened by a factor of 60 per cent inasmuch as the targeted decline in fertility rates from 1970 to 1976 requires 60 per cent more births to be prevented in 1970. The depressants are treated one by one:

(i) Program methods: IUD users are calculated as of the end of each calendar year by cohort projection. Continuation rates are optimistically assumed to improve for cohorts of 1972 and later, and no corrections are made for any shift to younger acceptors or for wasted wearing time, e.g., overlap with postpartum amenorrhea, secondary sterility, mortality, etc. Actually, most insertions now go simply

to replace annual terminations from the large body of wearers, i.e., the state of stability has been approximately reached. The plan as calculated, however, envisages nearly a 60 per cent rise in IUD wearers between 1970 and 1976 by an enlarged annual target of acceptors and by the assumption of an improvement in continuation.

Women protected by sterilization of self or spouse rise not by 60 per cent but 118 per cent, due to the introduction of a tubal ligation program in addition to a small rise in the annual vasectomy target. Under a cohort projection of the sterilization acceptors a discount of about 5 per cent is made each year to reflect secondary sterility, and no acceptors are counted past 10 years.

Pill and condom activity by the program is normally targeted in terms of users, that is, monthly numbers of couples to receive supplies. Pill users were assumed to rise by 88 per cent and condom users by 36 per cent over the 1970-76 period, after a discount of 10 per cent for wastage, maldistribution, etc.

Aggregating the four methods, a 67 per cent rise in numbers of couples protected results, somewhat above the 60 per cent overall requirement. Such a rise is thought feasible from the growth in numbers of married women before 1976, from deeper penetration of the group, and from the improved IUD continuation assumed. The percentage of couples above age 25 using program contraception would rise from 23 per cent to 33 per cent between 1970 and 1976.

(ii) Non-program contraception: Commercial sales of contraceptives in 1970 covered only about 5 per cent of couples; this was rising and a 60 per cent increase, to 8 per cent, could safely be assumed, especially as the component was a small one. This component implicitly incorporates all private birth control including non-program sterilization.

(iii) Abortion: The position taken is complex. The potential competition between attaining more abortion and more contraception is recognized. Stress on better IUD continuation is urged to relieve some of the need for abortion. A calculation is made that since it appears from survey data that abortion contributes less to fertility decline than program contraception, and since the latter will rise by 67 per cent rather than 60 per cent, then abortion need increase its contribution by only 48 per cent to produce a combined 60 per cent effect. (No action to this end is recommended, but the law bearing on abortion has since been liberalized.)

(iv) Later marriage: Fertility at age 15-19 was already negligible by 1970. At 20-24 it is targeted to decline more slowly in the future

than in the past, but the fall is still a substantial one. The plan assumes that later marriage will produce most of this, and birth control the rest, but it is unspecific as to the division. Throughout the document, all program-induced births prevented are considered to occur within the age groups above 25. This consumes the full force of the program, with nothing left over for ages below 25. The marriage age was already high by the end of 1971, and many observers doubted that it would rise much further. The large cohort of young women coming up may marry later, due to competition in finding mates who have the usual 4-5 years seniority in age. However, this cohort will not reach marriageable age until the late 1970s. What happens with the 20-24 age group during the 1972-1976 plan period is critical, as its size increases by a third from 1970-1976.

MOROCCO

The 1968-72 plan goal was to reduce the crude birth rate from 50 to 45 per thousand. That having failed, calculations were drawn up for two options—to reduce the CBR from 49 per thousand in 1972 to either 45 or 43 per thousand in 1977. Of these choices, the former, 45 per thousand, was selected (Royaume du Maroc) ([3]).

The plan considers only the program contribution, and assumes no effect from the private sector, no change in the age at marriage, and a constant mortality schedule. Changes in the age structure of the population are recognized only among females of childbearing age. These simplifications may cause error in the expected 1977 CBR, but the targeted *program* reduction in the CBR leads nevertheless to useful calculations on numbers of acceptors implied. The plan also illustrates how resourceful calculations can be devised in a situation of limited data.

The chain of calculations is fundamentally as follows. The 1971 census population is projected to 1977 at approximately 3.1 per cent annual growth, and the 1977 figure is multiplied by the reduction in the crude birth rate desired, to obtain the number of births to be prevented in 1977, which turns out to be 55,500.

To project the number of married women in each year, a distribution of females in five-year age groups is survived forward under fixed mortality conditions to mid-1972 and mid-1977. These two distributions are multiplied by the percentage married at each age

([3]) A typographical error appears in the last paragraph of page 855 of the plan, indicating that the goal of 43 was chosen. However, page 866 presents the acceptor targets which correspond to the 45 goal, and the unpublished documents are explicit in selecting 45. They are entitled "Impact Démographique" (see document, "Impact démographique et volume de travail").

(taken from 1960-62 data), with interpolation to obtain the yearly numbers between 1972 and 1977.

A set of age-specific fertility rates (ASFRs) is drawn from a large 1962 survey, and converted to MASFRs by the use of the same schedule of percentages married at each age. All three schedules are kept constant over time to represent the non-program projections.

The logic is that program acceptors will raise the number of users and so reduce the MASFRs, though unequally by age since younger women use contraceptives less reliably. In turn the ASFRs and the CBR will fall. Acceptors and users are assumed to have the same expected fertility as other women, and no allowance is made for timing within the birth interval or any of the related corrections.

To obtain the baseline (1972) program situation, IUD acceptors from 1966 on are treated in six-month cohorts and projected forward under life table continuation rates to December of 1972, giving 24,713 current users. (Continuation rates to 36 months are taken from a Tunisian local survey and after 36 months from Taiwan data). To these are added 10,000 current pill users, the 10,000 being the 1972 monthly average of women coming for pill resupply. Thus current users total 34,713 at the start of the plan. These users are distributed by age according to the age data for all program *acceptors* in 1970. (This same age distribution is used for all acceptors and users throughout the plan). Finally, users by age are multiplied by the 1972 MASFR schedule to yield births prevented in 1973, and these are subtracted from the total to produce the reduced 1973 MASFR schedule.

It is now easy to jump to 1977. From the original calculation showing the discrepancy between projected and desired total births in 1977, 55,500 births are to be prevented in 1977. These are distributed across age groups in the same percentage pattern as the results obtained when the baseline MASFRs are applied to 1,000 acceptors/users under the 1970 age distribution already mentioned. This gives the 1977 births prevented target by age, and the corresponding reduced 1977 MASFR schedule. Interpolation between 1973 and 1977 gives the annual MASFR target schedules, and these control the births prevented and acceptances needed in each year.

The question then is how to schedule annual acceptances to produce sufficient users and births prevented, with allowances for (a) dropouts and (b) the aforementioned unreliability of use by younger women. First the *targeted* births prevented at each age are calculated according to the difference between the targeted MASFR schedule and the non-program (1972) one. Next, the correction for unreliability of use is made by adjusting the potential fertility rate at each age. At

ages 20-24, for example, the reliability of use is only 45 per cent (4), so that the expected (non-program) MASFR of 376.6 at that age is multiplied by the 45 per cent and becomes only 170, considerably raising the number of young users needed to produce the births prevented as already calculated.

The actual number of users needed at each age in each year is simply the number of births prevented desired in that age group during the *next* year, divided by the MASFR just modified. For example, for ages 20-24, 5,750 births prevented are desired in 1975, and this is divided by the 170 rate to give 33,820 users needed in 1974. Such calculations generate the number of users by age by year.

Annual recruitment goals now follow easily. Enough new acceptors must be attracted to equal the calculated rise in users from each year to the next, a rise that, as will be recalled, flows directly from the interpolated decline in MASFRs from 1973 to 1977 and the growing number of births prevented required to produce them. (A contributing factor, of course, is the growth in the numbers of married women—a rise occurs in every age group in every year.) The number of acceptors desired is also augmented to cover an annual 15 per cent drop-out from the previous year's users, a factor applied equally to all ages. At last, by summation across all ages, the total number of acceptors needed for the year results.

A final note is in order on the relation between the 15 per cent annual termination rate among all users and the age-specific adjustment for unreliability of use. The two calculations say in effect that young women continue as well as older ones, but while continuing they experience on an average more failures with consequent childbirth. The 15 per cent calculation also says that the annual risk of termination is constant and that terminating women who later reenter the pool of acceptors experience the same risks as before.

On a 1973 population of about 12 million, acceptance targets for 1973-1977 are:

1973	70,220
1974	64,570
1975	74,300
1976	84,770
1977	97,530
Total	391,390

(4) The adjustments are taken from the results of the Taiwan KAP III Survey of 1970, and are 20, 45, 73, 88, 97, 99 and 100 per cent respectively for age-groups 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 and 45-49 (see Taiwan Provincial Department of Health, 1970).

PAKISTAN

The 1965-69 Pakistan program (Pakistan, Ministry of Health, Labour and Social Welfare) operated with an overall objective of reducing the crude birth rate from 50 to 40 per thousand, by reaching the 20 million fertile couples in the population. Assuming a concurrent decline in the crude death rate from 20 to 15 per thousand, the growth rate would fall from 3 to 2.5 per cent. The longer-range goal was a growth rate of 1.5 per cent by 1985 (Adil, 1968; Bean and Seltzer, 1968; Bean and Bhatti, 1969).

The dominant concern in target analysis was to translate program performance into couple-years of protection (CYP) against pregnancy, and from there to births prevented. The latter was accomplished by the simple rule that one birth was avoided for every three couple-years of protection. For the former the rules changed from time to time, and the following describes one of the chief variants.

During each month of program operation, IUDs, sterilizations, and conventional methods were accepted, and each of these was converted to the single measure of CYP. A sterilization was given a full CYP in the first year and 15 per cent less for each year thereafter. An IUD was assumed to give .75 CYP during the first year, .50 the second, .35 the third, and none thereafter. Conventionals were kept on a current basis: 100 units distributed in a particular month contributed 1 CYP, on the rough assumption of coitus twice a week.

All these assumptions were scientifically controversial, though their value for administrative simplicity was conceded. Objections included:

For conventionals: besides the uncertainty of the rule of coitus twice per week, no discount was made for supplies which were recorded as having moved through the distribution chain but failed actually to reach potential users, or for those which did so but were not indeed used, or for intermittent and defective use resulting in pregnancy.

For the IUD: The stair-step approach of .75, .50 and .35 CYP for the first three years called for improvement through use of a smooth decay function as widely employed in the life table method. Further, no explicit allowance was made for wasted wearing time through overlap with postpartum amenorrhea, secondary sterility, dissolution of marriage and other factors (for further discussion see section on appraisal of methods).

For sterilization (almost entirely vasectomies): In the first year, a vasectomy was said to give 1 CYP, which assumed that all couples remained alive and cohabiting, were fecund at the time of the opera-

tion and would have remained that way; that no wives were pregnant when the operation was done; that no post-operative failures occurred due to residual sperm; and that no women were amenorrheic due to a recent birth. For subsequent years, the 15 per cent discount per year was an effort to recognize such factors; whether 15 per cent was too much or too little was not determined.

Couple-years of protection for all three methods were summed, to produce a single index of achievement. This procedure made all months of protection equivalent, without differentiation as to whether the wife was amenorrheic, fecund or in a state of pregnancy during the particular month, and without regard to the decided advantage of staying with individual couples over an extended period of time in order to move them systematically into the fecund part of the birth interval and reduce the proportion of "protected" couples which was actually amenorrheic. But the overall procedure did have the appeal of giving a single measure of program achievement over time and between local areas. The hazard was that of invisible distortions of progress toward the desired fertility reduction, either by overstatement or indeed understatement of administrative accomplishment.

PHILIPPINES

The Philippine four-year economic development plan 1972-1975 (Republic of the Philippines, 1971) contains a calculation relating number of clinics, number of acceptors, and saturation of the eligible group, with the consequent effect on the population growth rate. The published work table is shown as Table 3.4.

The eligible group is estimated at 12.5 per cent of the population, and is considered to grow at 3.1 per cent per annum. To service this group, the program must develop a substantial level of activity, and this is made directly dependent upon the establishment of clinics. From the start in 1970, the number of clinics grows to 1,870 in 1974, and on the average each one recruits 44 new acceptors per month during 1971, declining to 25.8 per month in 1974. Simple multiplication of this number by the number of clinic months of operation within the year yields the acceptors registered for that year (Column 6). These acceptors are accumulated over time with a 4 per cent loss from each year to the next to recognize death and sterility. Cumulative acceptors grow much faster than does the target population, until in 1975 a large proportion of the then eligible group will have at some time accepted family planning in the clinics. No effect is included for private birth control or for later marriage. Census data do not indicate any large new cohort of young people entering the eligible group, and with the marriage age already high

Table 3.4. *Philippines: Targets for new acceptors at family planning clinics, FY 1972-1975 assuming 1,870 clinics by FY 1974 (based on number of clinics and remaining eligible women)*

Fiscal year	Number of eligible women at beginning of year (a) (in thousands)	Number of remaining eligible women at beginning of year (b) (in thousands)	Proportion of eligible women still remaining at beginning of year (c)	Number of acceptors per month per clinic (d)	Number of clinic months (e)	Number of registered acceptors during the year (f) (in thousands)	Cumulative number of acceptors at end of year (g) (in thousands)	Number of clinics to be added annually	Cumulative number of clinics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1970	—	—	—	—	—	100	100	—	—
1971	4,647	4,547	0.978	44.0	7,847	345	441	—	870
1972	4,791	4,350	0.908	40.9	14,076	576	999	400	1,270
1973	4,940	3,943	0.798	35.9	18,876	678	1,637	400	1,670
1974	5,093	3,456	0.679	30.6	21,876	669	2,240	200	1,870
1975	5,251	3,011	0.573	25.8	22,476	580	2,730	0	1,870

(a) 12.5 % of the population, assuming an annual growth rate of 3.1 %.
(b) Number of eligible women for current year minus cumulative number of acceptors.
(c) Ratio of remaining eligible women beginning of the year to eligible women at the beginning of the program.
(d) Proportion of women remaining multiplied by 45 (optimum average of acceptor rate/clinic).
(e) Allowance is given for new clinics. Number of old clinics multiplied by 12 mos. and number of new clinics by 9 mos.
(f) Estimated from clinic months and acceptors per month per clinic.
(g) Cumulative acceptors is adjusted for addition due to death and sterility by subtracting 4 % of previous year's cumulative figure before adding the current year's figure.

Source: Four-year Development Plan FY 72-75. (Adopted by the National Economic Council, 23 July 1971). Republic of the Philippines. Table 16.1 page 215.

Column 1 grows at 3.1 % per year.
2 = col. 1 minus col. 7 entry for previous year.
3 = col. 2 divided by col. 1.
4 = col. 3 times 45.
5 = see note (e).
6 = col. 4 times col. 5.
7 = col. 6 plus .96 times col. 7 entry for previous year.

for the region, the growth of the eligible population should be approximately the 3 per cent rate cited.

The plan emphasizes that the exact effect of this upon the population growth is unclear, as the calculations leave unspecified the continuation of use after acceptance. However, the plan refers to the possibility of converting the target of approximately 50,000 acceptors per month as of early 1973 to a number of "continuing users" and mentions that under high and low assumptions the program might bring the population growth rate from an estimated 3.1 per cent per year to a new level of from 2.3 to 2.7 per cent.

A separate document (Republic of the Philippines, 1972) carries all this forward, and uses a complex model to estimate births averted (Laing, 1972). Over the 1970-1976 period the age structure will change little, and non-program influences are considered minor. Hence it assumes that with no program the crude birth rate would stay at 43.2 and the crude death rate would fall by .3 points per year, for example from 13.2 to 12.9 per thousand between 1970 and 1971.

The methodology covers the same ground as in Table 3.4, though with somewhat modified figures. The same primacy is given to numbers of clinics, and these are now projected in three groups—two representing the two most active service organizations, one of which keeps the same number of clinics and acceptors from 1972 onwards, and one of which grows sharply due to the expected addition of 480 clinics. The third group represents all other organizations; its annual acceptances are made contingent partly upon the declining proportion of eligible women left unrecruited and so are dampened with time. Summation across the three groups of clinics provides the projection of total acceptance.

This projection however is reduced on two counts—once for over-reporting and transferees among clinics, and once for the same 4 per cent mortality and sterility factor as before. The corrected projection is then subtracted from all eligible women, as in Table 3.4., to give the number and proportion of eligible women left unrecruited at the start of each year.

The really new feature is the translation of acceptances to births averted. Without presenting the full details, the notable features are as follows: First, the method is founded upon a post-acceptance fertility expectation that disregards the duration of contraceptive use and any switching of methods. It takes the pragmatic view that, especially with sporadic-use methods such as rhythm, pills, conventionals, etc., the actual fertility record of acceptors is the most significant information to use. In the calculations, birth interval measures are used, on an age-specific basis and with separate data

(from a follow-up survey) by contraceptive method—pill, IUD, rhythm, other.

In one stroke this circumvents the entire apparatus of continuation rates, corrections for useless wearing time, projections of acceptor cohorts and so on. (After any post-acceptance birth, the risk of yet another birth, i.e., the expected birth interval, is assumed to revert to the same level as after acceptance.) Program effect is fertility registered after acceptance subtracted from the fertility expected, which is taken conservatively as equal to the prevailing age-specific marital rates in the general population. Numerous adjustments and corrections are made, partly to put all births prevented onto a calendar year basis and to account for peculiarities of the acceptor population. These are explained fully in the source, along with various refinements. The application of this method to the Philippine acceptor projections shows the crude birth rate falling from 43.1 to 36.9 per thousand and the growth rate falling from 2.99 to 2.55 per cent between 1970 and 1976. (Republic of the Philippines, 1972, Table 3.)

Note that the direction of calculation is from acceptances to births prevented, not the reverse; yet with the addition of assumptions about the age of acceptors and of distribution of methods, etc., the acceptances required for any fertility reduction target can be found, and a simplified formula is given for calculating the numbers of acceptors needed to reduce the crude birth rate by a given amount.

TAIWAN

The Taiwan plan for 1970-1979 (Taiwan Provincial Department of Health, 1970), contains two projections, one representing a weak program, termed the "high fertility" case, and one a strong program, the "low fertility" case. The plan runs from January 1970 through December 1979. First a population projection was prepared for the decade, incorporating a small decline in life table mortality from the 1969 levels. Fertility change was handled in two stages—first, that due to trends in proportions married and second, that due to trends in MASFRs.

In the marriage projections it is fortunate that only one age group, 20-24, really matters (Figure 3.1). It is large, it produces a substantial share of total births, and it is unpredictable in the future proportion marrying. To handle it (and all the other age groups as well) the principle followed was to project graphically the 1960 to 1969 trend line. The 15-19 age group already has a very low proportion married, and the groups over 25 have very high proportions, which are conservatively assumed to rise yet further except for the 25-29

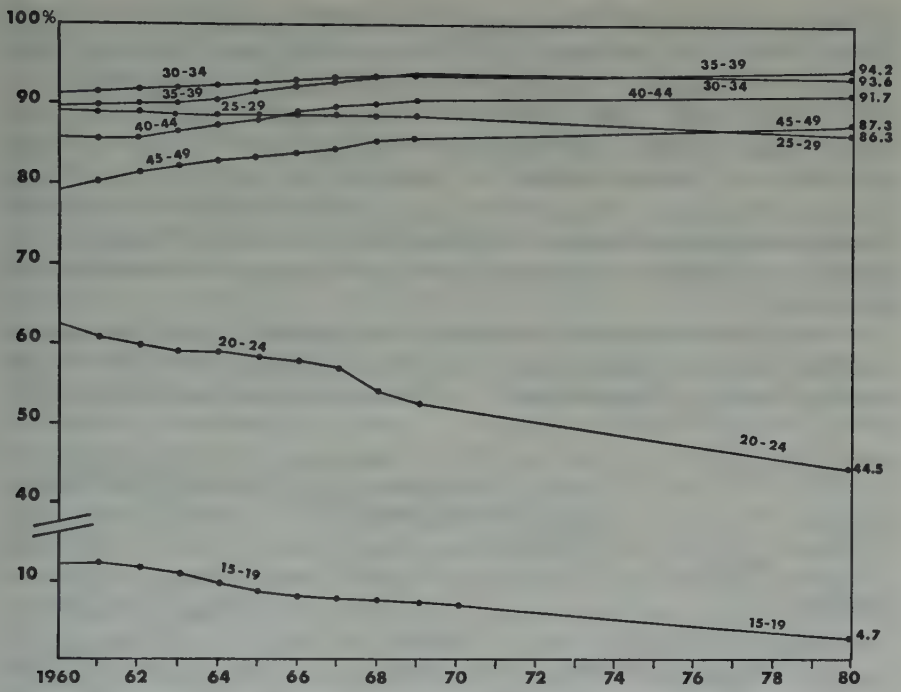


Fig. 3.1. *Taiwan: Actual and projected percentage of women currently married, 1960-1980*

group, whose proportion declines very slightly from near 90 per cent to 86 percent over the decade. Thus reasonably safe projections could be made for all age groups but one. The annual proportions married shown in Figure 3.1 were applied to the basic population projection to project the annual numbers of married women.

The next step was to project the MASFRs (Figure 3.2), once for the "high fertility" and once for the "low fertility" case. The former left the rates for the 15-19 and 20-24 age groups at their high 1969 levels, and for all other groups simply projected the downward trend line of the 1960-1969 period. For the "low fertility" case, the 15-19 and 20-24 rates, which had risen in the 1960s, were made to decline back to their 1959 level and for ages 25-49 each MASFR was declined linearly to the Japanese 1950-55 level, with some modifications.

As Figure 3.2 shows, the divergence between the two projections clearly matters little for the age groups from 35 to 49. And for the 15-19 group so few women are married that the divergence in rates, though large, also matters little. It is the three groups for 20-34 that are determinative.

The logic for the calculations then flows as follows: If there were

no birth control at all in the society the MASFR would be at very high levels (which are estimated partly from the highest historical rates recorded in Taiwan). To bring these rates down to the schedule for the "high fertility" case requires a large proportion of women practicing contraception and indeed in the base year of 1970 a large proportion of Taiwan women do practice (due to both private and program influence). This is taken as a given, and all future *increases* occur through the program. To attain the "low fertility" case everything is the same except that stronger acceptor increments are required.

Thus three schedules of MASFR are applied to the numbers of married women in each year: (a) "contraception-free", (b) "high fertility" and (c) "low fertility". The a-b difference and the a-c dif-

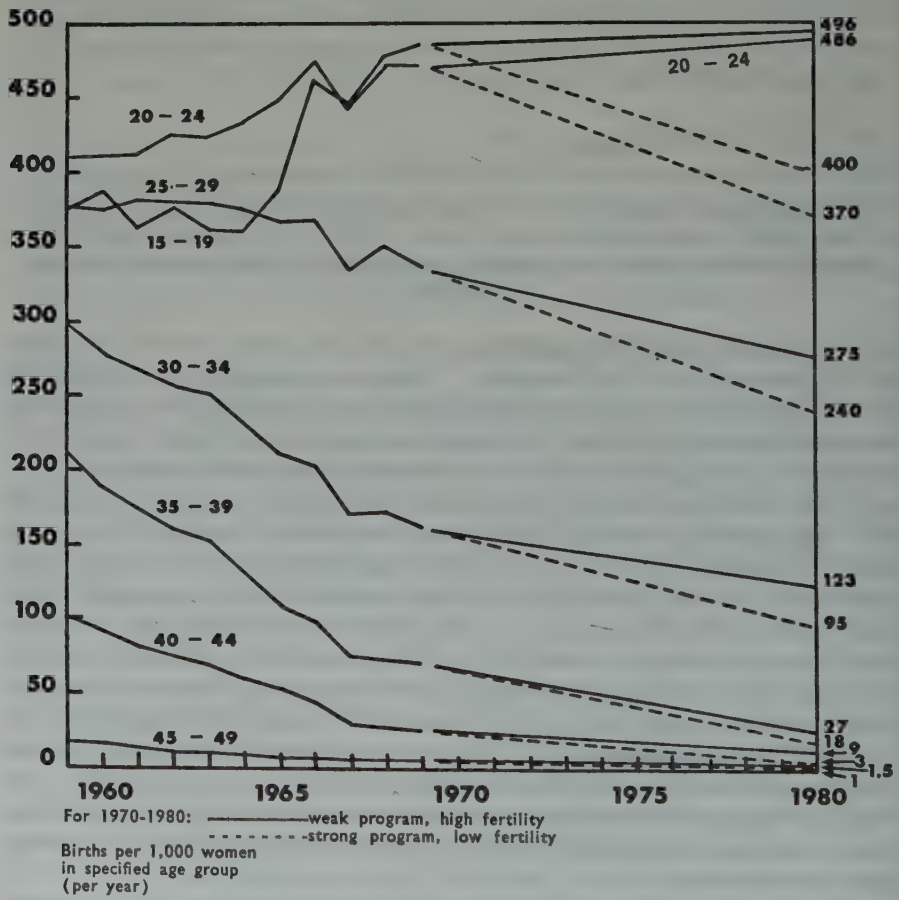


Fig. 3.2. *Taiwan: Actual and projected age-specific marital fertility rates, 1959-1980*

ference give the numbers of births to be prevented under the two levels of program strength. In neither case does the program need to do the whole job outright—it starts from the spring-board of the 1969 practice level, and must only add on the acceptor increments necessary to reduce the MASFR a little more each year. The chief interest of course lies in the b-c difference, which is fortunately independent of the (a) level.

The births to be prevented are translated to numbers of practicing women by a 4-step method. First, acceptors are assigned a (high) age-specific schedule of potential fertility taken from studies of IUD acceptors' pre-acceptance fertility. Second, this schedule is reduced by an index of unreliability of use, developed to reflect the rather high fertility actually registered by women identifying themselves in surveys as contraceptive users. (This is the index borrowed by the Moroccan plan and discussed in that section. It reduces program effect upon potential fertility most for acceptors aged 15-19 and less with each older group, the multipliers being .20, .45, .73, .88, .97, .99 and 1.00 across the seven 5-year age groups from 15-49.) The result is a modified schedule of births prevented in the next year to produce the required number of users in this year. The rise in users over last year gives the acceptor recruitment target, except for an addition equal to 15 per cent of last year's users, who are assumed to have terminated. (This feature was also borrowed in the Moroccan plan and is discussed there.)

Private contraception is reflected in the 1969 base through the percentage already practicing, but abortion is not, and neither factor is recognized in the annual gains from 1970-79. The 15 per cent termination rule is simple but undefended; any program experiences a constant dropout after several years of approximate stability in annual acceptances and termination rates, but if the average wearing time is, say, three years, then 33 per cent would drop out annually, and more if wearing time is less. No method mix is used and, as with most plans, calculations are not made on the cumulative proportion of each age cohort ever accepting. There is a special problem with the 15-19 married age group since 86 per cent of it must be using contraception in 1979 in the low-fertility projection, apparently reflecting the dual impact of a large drop in marital fertility (Figure 3.2) and the low reliability of use factor (.20). The 20-24 group suffers from the same problems but is partly relieved by a higher (.45) reliability of use.

Due to a great increase in the 1970s in the size of the already large 25-29 age group, together with its constantly high proportions married, a near explosion in its number of current users is required. Under the low fertility assumption the annual increment in users goes

from about 14,000 per year in 1970-72 to above 44,000 in 1979, bringing the total users to about 390,000 in 1979 or 58 per cent of the group. The high achievement levels expected here and elsewhere in the plan are paralleled by impressive results; under the stronger program the proportion of married women aged 15-49 currently using contraception rises from about 45 per cent to 66 per cent during the decade and the crude birth rate falls from 26.6 to 22.5 per thousand, despite the worsening age structure.

Overall it turns out that about 4.13 users over one year prevent one birth, equivalent to 242 births prevented per 1000 user-years.

The Taiwan plan goes beyond the generation of acceptor targets to produce cost figures for the two alternative programs. These are interesting in both method and results and are available from the source document.

THAILAND

Thailand's goal is to reduce her rate of natural increase from the 1970 level of something over 3 per cent to 2.5 per cent by 1976 (Unhanand *et al.*, 1972; Thailand, Ministry of Public Health, 1971). Calculations have been made indicating that the program can reduce the crude birth rate by about 8 points (from 41.0 to 33.2) by the end of 1976, using 1970 as a base. The reasoning runs as follows:

The three contraceptive methods, pill, IUD and female sterilization, are well established in the program and are expected to grow in popularity. The pill particularly may enjoy wider usage, as it is dispensed by medical auxiliaries who are widely dispersed in the rural health facilities. The Thai plan gives a year-by-year schedule of the numbers of acceptors required for each of these three methods, with the total increasing from 300,000 for 1971 to 410,000 for 1975 and 1976. This is an annual acceptance figure of 5 to 8 per cent of the married women aged 15-44. It is recognized in principle that some women will try more than one method and will be duplicated in the counts, and that some acceptors will pass out of the eligible group (by death, broken marriage or aging), so that as of 1976 an appreciable proportion of the past acceptors must be discounted. However, none of these considerations enters the calculations.

How many births will these acceptances prevent? A smooth decay curve of continuation is used for acceptors of the IUD and the pill, although at high continuation levels, especially for the pill. For sterilization acceptors no decay is assumed. Using time is distributed in the form of couple-years of protection, with a nine month lag to births averted.

As in Korea, past acceptors play a significant role—1965-1970

acceptors contribue 35 per cent of the couple-years of protection present during the 1971-1976 period. It is unclear how much their greater aging would require larger discounts in useful wearing time.

The conversion to births prevented is straightforward: 3.25 couple-years of protection are considered to represent one birth prevented, equivalent to the assumption that 3.25 years is the normal birth interval for acceptors. Two adjustments underlie this figure. The normal Thai birth interval well exceeds three years, but acceptors are usually of higher fertility, so 3.0 is considered reasonable. On the other hand, acceptors often start use soon after a birth when they are still amenorrheic and so "waste" using time; therefore 3.25 is better. This necessary guesswork results in an assumed fertility-prevented rate of 1/3.25, or 308 births per 1000 woman-years of protection.

To estimate the effect on the crude birth rate, the official population projections for Thailand (low-growth assumption) were taken as a baseline. The number of births prevented for each calendar year was put against the estimated population for that year to yield the change in the crude birth rate, and the calculated result is, as mentioned, a very large reduction in fertility. So far as the course of the crude birth rate is concerned, it will probably be driven down by the program less than calculated, but will be strongly affected by the marriage age and by private birth control.

TUNISIA

Tunisia has developed a most interesting plan (République Tunisienne, 1972), embracing many important features. Its goal is to achieve the same gross reproduction rate (GRR) and individual age-specific fertility rate (ASFR) by 2001 as Italy had in 1970, and to move toward them by a linear path at each age. By 1981 ten years of the thirty-year span will have passed, and one-third of the fertility gap should have been closed (Figure 3.3).

Note that this progress would reflect marriage changes and private birth control as well as program effect. And, of course, the growth rate and crude birth rate might still fluctuate depending on the age distribution and on mortality.

The plan calculates the number of births *desired* in 1981 by multiplying the 1981 curve (Figure 3.3) by the age distribution of females projected for 1981. It calculates the number of births *expected* for 1981 by simple extrapolation to 1981 of the 1966-1971 trend in each marital age-specific fertility rate, multiplied by a set of 1981 predictions of the number of females married at each age (Marcoux). The 1966-1971 trend is mildly downward and this, pro-

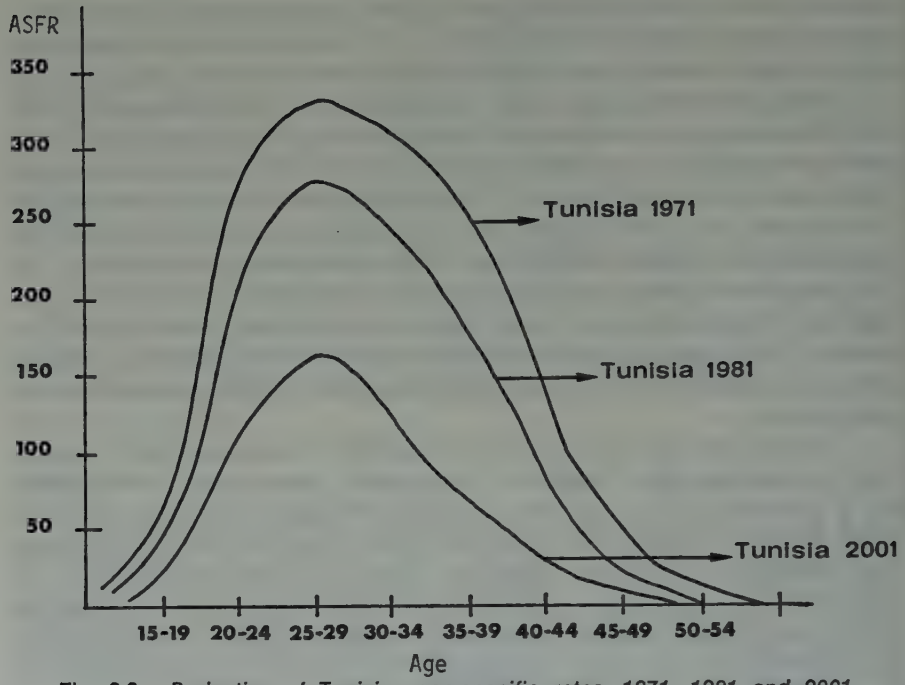


Fig. 3.3. *Projection of Tunisian age-specific rates, 1971, 1981 and 2001*

jected to 1981, is taken to represent secular tendencies in private birth control. Of course, the program also operated in the period 1966-1971, but the estimated births it prevented in this period are added back in before the extrapolation to 1981 is made.

The difference in the two figures, expected versus desired, is the number of births to be prevented in 1981. Annual targets rise steadily from 12,500 prevented births estimated for 1971, to 49,500 by the year 1981, with larger annual increases toward the end of the period. Small additions are made for 1974, 1975, and 1976 to cancel earlier shortfalls.

How then to make each year's target? In two ways: by users of contraception/sterilization, and by abortions. Users are calculated from past acceptances, recognizing the substantial drop-out rates which permit only a modest rise in users despite numerous acceptances (about 25,000 new acceptors in 1973 add only 5,700 users in 1974). Then only one birth is considered to be prevented next year for every four users this year (regardless of method, whether IUD, pill, condom, jelly, or sterilization). Age and other sub-group distributions are ignored in the acceptance-user calculations, despite the clear goal of a linear fertility decline at each age. For women

who abort, six births are prevented for every ten abortions. Under the schedule of past and projected acceptances, it works out that 40 per cent of the 1974 target must come from abortion and 60 per cent from the 1973 users of contraception/sterilization.

TURKEY

Turkey's population growth goals have changed frequently over the last decade, but a plan published recently by the Ministry of Health (Turkey, 1972) is interesting for its manner of handling the target question.

The plan starts by reviewing Turkey's past and projected population growth, together with information on various fertility measures. A carefully done projection to the year 2000 is invoked, which assumes that age-specific fertility and mortality rates will continue to fall gently, at the same pace as in the recent past. This reflects chiefly the secular trend in age at marriage, health practices, and birth control, as the program itself has so far been a limited one with no significant effect upon the birth rate. By the year 1995 the population size would reach 68 million, a level seriously above the State Planning Organization's published aim of no more than about 60 million by that date.

From this discrepancy the plan develops its objectives. To meet the 60 million target, it estimates that the total fertility rate (TFR) will have to fall from 5.7 in 1972 to 3.0 in 1995. For the first five years, mid-1972 to mid-1977, the plan set a target of a 0.5 drop in the TFR.

With attention narrowed to those five years, the female population (in five-year age groups) is moved forward year by year, and an unchanging set of age-specific fertility rates is applied to obtain a baseline number of births anticipated each year. The process is then repeated under three alternative levels of program strength through acceptances set at low, medium, and high rates. Each five-year age group starts in 1972 as a cohort and moves forward through the acceptance experience of each assumption. A smooth decay curve for continuation is applied (with very favorable assumptions about duration of use), the same curve being used for each of the three levels of program acceptance, to a number of continuing users as of the mid-point of each year, who are removed from the risk of having live births. The number of births expected at each age, i.e., the baseline mentioned above, is reduced by an exactly proportionate amount. This changes the age-specific fertility rates, with a modified TFR and other fertility measures directly following. The result is that the TFR would fall as targeted from 5.7 to 5.2 under the third assumption, of low acceptance rates. However, the continuation rates

assumed are extremely high ones and the age distribution of acceptors is optimistic with respect to the participation of younger women, whose fertility is high. On the other hand, the calculations ignore any assistance from a rising age at marriage, or from increased abortion and contraception in the private sector.

One feature of the plan deserving special mention is that each acceptance age cohort is carried through time separately, and is subjected to a rate of increasing acceptance figures on the base of *current non-users.* This avoids the pitfall of across-the-board acceptance rates which if attained year after year would sometimes accumulate more users in certain age groups than there are members. A difficulty remains, however, that with time a large proportion of a cohort may discontinue use of the contraceptive methods offered, making the assumed acceptance rate among non-users perhaps unrealistic.

This plan has the strength of component projection both for the population and for acceptor experience. It is structured within the context of careful baseline information from previous demographic work, with estimates of age-specific fertility rates, an age distribution, and a reasonable population projection. The work incorporates several fertility measures not just one, and shows how progress in reducing the TFR is diluted by an unfortunate age distribution so that the crude birth rate and population growth rate are not able to fall as impressively as the TFR. However, no attention is given to states of pregnancy, amenorrhea, secondary sterility, etc. Nor is there recognition that acceptors may have higher fertility expectations than others, with therefore more births prevented than the prevailing age-specific fertility rates would indicate.

Unfortunately, it happens that population growth will be especially high within the age range 20-34 where the highest fertility rate prevails. The high rate aggravated by the growing base creates a powerful pressure toward growth even in the presence of some program gains.

III. APPRAISAL OF METHODS USED

How can these fourteen plans be usefully contrasted and assessed? Clearly, the data available in each country, and that used in the plans, varies widely. Goals are stated in different terms, the birth control methods used differ, and the structures of reasoning and calculation vary. Table 3.5 summarizes selected features of the fourteen plans, and the following paragraphs take up the rows in order.

Ten plans state their demographic aim quite explicitly; four are

couched in softer terms but present calculations linking demographic effects and acceptance. The plans agree remarkably in targeting a fall of about one point per year in the CBR; two exceptions are Taiwan at half that but starting from the low level of 27 (see text), and Pakistan at double that but starting from the high level of 50. One point reduction per year is, on the basis of past experience, very optimistic for program effect alone.

Multiple methods of contraception are used in all the twelve plans giving information on the question. (The Taiwan document is unspecific, but the IUD, pill, and condom are used.) Seven plans make no mention of sterilization and only three use both male and female sterilization. All use the IUD, almost all the pill, nine either the condom, conventionals, or other. The Philippines plan includes rhythm as a major method. Program abortion is part of the Tunisia plan (and has been added on a small scale recently in Korea).

To relate user-years to births prevented, nine plans use a simple rule of 1:3 to 1:4. No clear distinction is made between one user for a full year and several users summing to a year. The former dominates, of course, where there are more sterilizations, higher continuation rates, and a decreasing weight of recent acceptors as the program ages.

Surprisingly, six of eight plans assume acceptors' potential fertility to be equal to that of other married women, not higher. This position is conservative, both in raising the numbers of acceptors needed and reducing the estimate of program effect, but it partially offsets other features which err in the opposite direction.

Turning to non-program influences, only three plans consider marriage and only four private birth control, whether contraception, sterilization or abortion. Since trends for these determinants are almost everywhere tending to help the programs, their neglect is also conservative. Age distributions are used in eight plans, to take account of the increasing numbers of women in the reproductive ages caused either by general population growth or by outsized young cohorts moving up.

The mortality factor is especially important to plans whose goal is stated in terms of growth as distinct from fertility. Of the seven plans considering mortality, India, Taiwan, and Turkey do so by a full life table as part of a component population projection.

Acceptors are treated by age in only six cases, and only one plan, Turkey, follows cumulative acceptances in each age cohort to watch for saturation. The omission of acceptors' age yields simplifications in the calculations, but supplementary work at least should be done to keep abreast of the patterns by age and living children.

Table 3.5. *Plan features by country*

	Ceylon	Ghana	India	Indonesia	Jamaica	Korea	
Aim	*reduce CBR from 36 to 25*	*reduce CBR 1 pt/year*	*reduce NIR of 2.5 by half*			*reduce NIR from 2.9 to 2.0*	*reduce NIR to 1.5*
Plan Period	1966-1975	1970-74	1968-9 to 1978-9	1971-75	1969-72	1962-71	1972-76
Methods	IUD, pill, condom	unspecified	st, IUD, conv	IUD, pill, conv	unspecified	vas, IUD, other	st, IUD, pill, condom
Birth Prevented per User-year	0.250	0.250	0.250 for conv	0.333	0.250 to 0.333	0.250	
Potential Fertility Acceptors/Others	equal		equal			equal	
Marriage							partial
Private Birth Control	X					X	partial
Age Dist. Projected			X			X	X
Mortality			X			X	X
Acceptors Treated by Age			X				
Saturation of Age Cohorts							
Acceptors Projected by Cohort for IUD and/or Pill			X	X	X	X	X
Comment on Continuation	Ignored; post-acceptance fertility is used	each year 1/3 of all users drop out					
Discounts for Ineffective Use	partial		only mortality for vas & IUD			partial	partial
Prevented Births Dist. by Calendar or Fiscal Year		X	X	X	X	X	X
Est. of Effect on CBR	before-after	change	before-after	change	change	before-after	before-after
Annual Recruitment Goal	5.3 % of MWRA	5 to 14 per 1,000 pop/yr	17 to 35 per 1,000 pop/yr	2-14 % of MWRA (15-44)	10 to 16 per 1,000 pop/yr	Up to 10 % of MWRA 20-44 for IUD and vas	About 8 % of MWRA (15-49) for st and IUD.

	Morocco	Pakistan	Philippines	Taiwan	Thailand	Tunisia	Turkey
Aim	*reduce CBR from 49 to 45*	*reduce CBR from 50 to 40, NIR from 3.0 to 2.5*			*reduce NIR from > 3 % to 2.5 %*	*reduce GRR from 3.11 to 2.46*	*reduce TFR from 5.7 to 5.2*
Plan Period	1972-77	1965-69	1970-76	1970-79	1970-76	1971-81	1972-77
Methods	IUD, pill	st, IUD, conv	IUD, pill, rhythm, other	unspecified	TL, IUD, pill	TL, IUD, pill, conv, abortion	IUD, pill
Birth Prevented per User-year		0.333			0.308	0.250	
Potential Fertility Acceptors/Others	equal		equal	higher	higher		equal
Marriage				X		X	
Private Birth Control						X	
Age Dist. Projected	partial		partial	X		partial	X
Mortality	partial		X	X			X
Acceptors Treated by Age	X		X	X		partial	X
Saturation of Age Cohorts			partial				X
Acceptors Projected by Cohort for IUD and/or Pill					X	X	X
Comment on Continuation	each year 15 % of users drop out	stairstep decay for CYP from IUD and st	Ignored. post-acceptance fertility is used	each year 15 % of users drop out			
Discounts for Ineffective Use			partial (4 % annual discount of users.)		partial		
Prevented Births Dist. by Calendar or Fiscal Year	X	X	X	X	X	X	X
Est. of Effect on CBR	before-after	before-after	before-after	before-after	change		before-after
Annual Recruitment Goal	5-8 per 1,000 pop/yr		2-10 % of MWRA	10-13 % of MWRA (15-49)	5-8 % of MWRA (15-44)	About 6 per 1,000 pop/yr plus abortions	3 levels

CBR: crude birth rate.
GRR: gross reproduction rate.
NIR: natural increase rate.
TFR: total fertility rate.

st: both male and female sterilization.
vas: vasectomy only.
TL: female sterilization only.
conv: conventional contraceptives

Continuation of use is the chief culprit in over-estimating program effect. The plans not only tend to adopt excessively favorable rates, they also fail to correct them sufficiently for ineffective use time. No plan considers the timing of acceptance, to balance acceptances soon after a birth, which overlap with postpartum amenorrhea, against those delayed after a birth, which lose more of the especially fecund women to another conception.

The corrections to using time deserve a further comment. The life table projection of wearing time is a mathematical construct. As mentioned at several points above, some "protection time" becomes useless by the death of one partner, by marital dissolutions of other types, by secondary sterility present at insertion or developing thereafter, by amenorrhea during the early months of wearing due to a recent birth, and by pregnancies *in situ.* This list of factors, in Potter's Taichung analysis (1969), caused a discount of 31 per cent (medium assumption) of the useful protection time projected under the life table curve. This is a most serious result, as many plans make the calculated fertility effect directly proportional to use time under the life table curve, and a nearly one-third reduction in effect on the birth rate would force fundamental revisions. The correct adjustment will vary by country, of course, but it has so far received inadequate attention. [Potter (1972) has recently refined his discussion of the various corrections.]

Virtually all plans manage to distribute prevented births by calendar time, to fit them into their planning cycle of fiscal years. None pursue the timing factor to consider the possible bunching of births to acceptors soon after termination due to their concentration at that point in the fecund state. However, such bunching may be neglected if a subtraction has been made for use during postpartum amenorrhea, and if the program is old enough to approach an even flow in numbers discontinuing contraception.

The Ceylon plan and the newer Philippines calculations go directly from acceptance to subsequent fertility experience, and if data are collected on a sufficiently long post-acceptance period, this embraces the occurrence of post-termination births on whatever displacement schedule they appear. Post-termination fecundity status, countervailing birth control practice, etc., are automatically incorporated, and though this approach does not completely solve the problem of precise time allocation of prevented births, it does keep acceptors and their fertility incorporated in the calculations regardless of discontinuation.

The procedure has a few drawbacks—that it is inattentive to the means by which post-acceptance and post-termination fertility are controlled, that it is unstable if some of these means (abortion,

private sector) change swiftly, and that it diverts attention from the need of many acceptors for 10-15 years of reliable protection when in fact they will receive only 1-3 years from the program's methods.

Four of thirteen plans simplify by estimating only the reduction in the crude birth rate, not the precise before and after levels. Thailand, Indonesia, Jamaica, and Ghana put births prevented by the program against the approximate population base as a denominator, to show the probable CBR reduction. This procedure relieves the calculations from having to specify what the rate is either before or after the action period, although for illustrative purposes approximate levels may be given. Components in the non-program sector are ignored.

Finally, the plans vary considerably in their recruitment goals. Six are clustered at about 5-10 per cent of married women of reproductive age per year, although Korea adds large numbers of pill and condom users. Five range from 5 to 16 acceptors per thousand population per year; India's plan uses figures rising from 17 to 35.

Turning from Table 3.5 to other features, the Tunisia and Turkey plans, although different from each other in many respects, both couch their aims in terms of age-independent fertility rates, the GRR and TFR respectively. They do not lose sight of the cruder rates, upon which political attention at least will continue to rest, but they narrow in on finer measures closer to program effect. The ones they choose also reflect marriage trends and private sector influences; Tunisia includes a separate projection for proportion married at each age.

As noted, Turkey's is the only plan to handle each age group as a separate cohort, keeping track of the cumulative percentage who ever accept.

No plan considers alternative *paths* of fertility decline. It would, for example, produce long-run gains if more of the reduction came early in the plan period; or it might be reasoned that a late reduction would give needed time to build program infrastructure. Most plans graduate the rate of acceptances to higher levels through time; the controlling factors appear to include how rapidly more budget can be gained, how quickly public interest will grow, how fast staff and services can be extended, etc.

The 1972-76 plan for Korea uses a proportionate device to project future program strength against future fertility decline, namely that a 60 per cent rise in numbers of couples protected by program and non-program determinants will produce a 60 per cent further rise in numbers of births prevented. This greatly simplifies, but it assumes that the relation of births prevented to program activity is a linear

one. Actually, the relation may either deteriorate or improve as the program matures. If sterilizations rise as planned and are accepted by more younger couples, if the IUD continues to be available only through the program sector and is used by younger women, and if program contribution increases through a helping (interaction) effect from modernization, the relation may sharpen ([5]). It may, however, deteriorate if a saturation point is reached or if the private sector grows strongly and undercuts the need for a special government program.

None of the plans recognizes the dangers inherent in the gradual disappearance of lactation. With migration to cities, greater employment of young women, and general modernization, breastfeeding diminishes, perhaps rapidly. Watson (1973) has estimated for Korea that a decline from 95 per cent to 85 per cent in the proportion of delivering women who lactate, accompanied by a decline in mean lactation time from 12 months to 10 months, would lose as much protection time as is afforded by the Korean national pill program, and a good deal more than the condom or sterilization program. For Taiwan, A. K. Jain and associates (Jain *et al.*, 1970; Jain and Sun, 1972) credit lactation with a delay of 7 months in the return of menstruation, both for all married women 20-39 and for IUD acceptors. If the mean birth interval is taken as 30 months, lactation thus explains 23 per cent of it. Partially compensating however for any decline in breastfeeding is a reduction in infant mortality, as the survival of infants lengthens breastfeeding. How great these opposing effects may be under different circumstances is not clear.

Much more analysis could be done of these plans than has been presented here. Nothing has been said here about the internal allocation of targets to local areas, nor about the staffing and budgetary components that some plans include. Numerous scientific points invite further attention, as for example the extent to which couples, when services and encouragement are firmly established in the environment, make micro-adaptations to their own perceived risk of conception and so confound the plans' use of averages. All this points to the need for continued exchange between these plans, developed as they are from administrative necessity, and the deeper analytic work that may improve their accuracy.

([5]) Berelson (1974) has tabulated 26 family planning programs, showing that at a given level of program strength, response is better in countries with more favorable institutional settings. The measure of response is 1971 program acceptors as a percentage of MWRA non-users.

REFERENCES

Adil, Enver

1968 "Progress and problems of fertility control around the world", Demography, Vol. 5, No. 2, page 660.

Bean, Lee L. and A. D. Bhatti

1969 "Three years of Pakistan's new national family planning program", The Pakistan Development Review, Vol. 9, No. 1, (Spring).

Bean, Lee L. and William Seltzer

1968 "Couple years of protection and births prevented: A methodological examination", Demography, Vol. 5, No. 2.

Berelson, B.

1974 "An evaluation of the effects of population control programs", Studies in Family Planning, Vol. 5, No. 1 (January).

Chandrasekaran, C. *et al.*

1971 "Some problems in determining the number of acceptors needed in a family planning programme to achieve a specified reduction in the birth rate", Population Studies, Volume XXV, No. 2 (July).

Gaisie, S. K. and S. B. Jones

1970 "Ghana", Country Profiles, The Population Council (October).

Henry, Louis

1953 Fécondité des Mariages, Institut National d'Etudes Démographiques, Travaux et documents, Cahier No. 16, Presses Universitaires de France.

1961a "La fécondité naturelle: observation-théorie-résultats", Population, 16, pp. 625-632.

1961b "Some data on natural fertility", Eugenics Quarterly, 8 (June).

International Bank for Reconstruction and Development (IBRD)

1972 Jamaica National Family Planning Program, Report of First External Review, Report submitted to the Government of Jamaica, Washington D.C., Annex 6.

Jain, Anrudh K. and T. H. Sun

1972 "Inter-relationships between socio-demographic factors, lactation and post-partum amenorrhea", Demography India, (Journal of the Indian Association for the Study of Population), Vol. 1, No. 1 (October).

Jain, Anrudh K. *et al.*

1970 "Demographic aspects of lactation and post-partum amenorrhea", Demography, Volume 7, No. 2 (May).

Jain, Sagar C.

1971 "Comparative study of effective and non-effective family planning programs in India", April 6 (mimeo), Department of Family Planning, Government of India, Appendix A.

Laing, John E.

1972 Births Averted by a Multi-Method Family Planning Program, Family Planning Evaluation Office, Population Institute, University of the Philippines (submitted for publication).

Marcoux, Alain

"Sur les facteurs de l'évolution passée et future des naissances en Tunisie", undated.

Pakistan, Ministry of Health, Labor and Social Welfare

Family Planning Scheme for Pakistan: During the Third Five-Year Plan, 1965-1970. Rawalpindi.

Perera, T. *et al.*

1968 "Targets for Ceylon's national family planning programs in 1970", Report of the Evaluation Sub-Committee to the Family Planning Advisory Committee, documents of July 26 and September 26, 1968 and November 13, 1969 (unpublished).

Potter, Robert G.

1969 "Estimating births averted in a family planning program", in S. J. Behrman *et al.* (eds.), Fertility and Family Planning, Ann Arbor, University of Michigan Press. See p. 425, Table 3.

1972 "Contraceptive impact over several generations", in T. N. E. Greville, Population Dynamics, Academic Press Inc., New York and London.

Republic of the Philippines

Four-Year Development Plan, Fiscal Years 1972-1975, pp. 214-215.

Republic of the Philippines, Commission on Population

1972 The Five-Year Population Program for the Philippines, Annex No. 1 (Mimeo).

République Tunisienne, Ministère de la Santé Publique

1972 "IVe Plan, 1973-1976. Objectifs Globaux du Planning Familial", Institut National du Planning Familial et de la Prévention Maternelle et Infantile.

Ross, John A.

1970 "Post-partum program: Action implications" in G. I. Zatuchni (Ed.), Post-partum Family Planning, McGraw-Hill, Table IV.

Royaume du Maroc, Secrétariat d'Etat au Plan, au Dévelopement Régional et à la Formation des Cadres

Plan de Développement Economique et Social, 1973-1977, Vol. II, Développement Sectoriel, Morocco, Undated.

Taek Il Kim

1970 "The Ten-Year Plan, 1962-1971" in Population and Family Planning in the Republic of Korea, Vol. I, Ministry of Health and Social Affairs, Republic of Korea.

Taek Il Kim *et al.*

1972 The Korean National Family Planning Program, New York, The Population Council.

Taiwan Provincial Department of Health, Committee on Family Planning

1970 Projection of Population and Corresponding Family Planning Recruitment and Costs by 5-Year Age Groups, 1970-1979 for Taiwan Area.

Thailand, Ministry of Public Health

1971 Targets for Number of Acceptors in Thailand's National Family Planning Program and Birth Rate Reductions Expected, 1971-76, Annex I, Population

Sector, National Economic Development Board and National Family Planning Program.

Thailand

Statistical Yearbook

Turkey, Ministry of Health and Social Welfare

1972 Population Planning Strategy, Targets, Outline of Programs and Inputs Needed for the 1972-1977 Five Year Period, A policy statement by the Ministry of Health and Social Welfare, Ankara, March 20.

Unhanand *et al.*

1972 "Thailand", Country Profiles, The Population Council, New York (March).

United Nations *et al.*

1970 Indonesia - Report of a UN-WHO-IBRD Family Planning Mission, Report No. 2, Ex. 2, April, pp. 45-54.

Watson, W. B.

1973 "Example of demographic significance of declining lactation" (mimeo).

Watson, W. and D. Smith

1971 "Determination of targets for the Korean Family Planning Program; Third five-year plan: 1972-1976," Developed for the Ministry of Health and Social Affairs, Republic of Korea (mimeo).

Wright, N. W.

1974 Personal communication, April 2.

Chapter 4

BIOSTATISTICS OF HUMAN REPRODUCTION

by
H. Leridon

I. INTRODUCTION: MICRODEMOGRAPHIC AND MACRODEMOGRAPHIC APPROACHES

In the analysis of the fertility of acceptors and non-acceptors the fact that the "fecund state" is neither permanent nor homogeneous throughout a woman's reproductive period cannot be ignored. In practice a classic problem of defining the population "at risk" is raised.

Take, as a simple example, the case of a post-partum program in which the potential acceptors are women who have just delivered. A fertility rate calculated for the year preceding the date of acceptance would closely approximate 1,000 per thousand. In contrast, the fertility rate for the first year after acceptance would be near zero, if only because of the time lapse between the date of conception and the delivery date. Even in less simplistic situations, it is clearly possible that the end of the period of post-partum sterility will overlap the beginning of the acceptance period.

A "microdemographic" analysis of fertility is a necessary first step in overcoming these problems. The biological aspects of fertility which we will discuss will serve as a background for some of the analyses and methodology set forth in this volume.

We will begin by reviewing the accepted demographic characteristics of natural fertility, a concept to which we will return later; let us briefly define a "system of natural fertility" as one in which neither of the individuals (nor the couple jointly) consciously attempts to regulate the number or spacing of their children.

After presenting these data on the "macrodemographic" aspects of fertility, we will turn to the microdemographic approach and an analysis of the components of natural fertility. These discussions will provide a framework into which contraception can easily be fitted.

Table 4.1. *Fertility in selected non-contracepting populations*

Population	Ref.	Marital fertility rate (per 1,000)						Mean number of children	
		20-24 years	25-29 years	30-34 years	35-39 years	40-44 years	45-49 years	Completed families, women married at the age of 20	All women surviving to age 45
Hutterites (Marriages 1921-30)	1	550	502	447	406	222	61	10.9	9.5 (a)
Canada (Marriages 1700-30)	2	509	495	484	410	231	30	10.8	5.65 (a)
Norway (Marriages 1874-76)	3	396	380	341	289	180	41	8.1	—
Martinique (birth cohort 1914-28)	4	481	440	333	211	114	11	7.9	5.4
Punjab Chamars (11 villages) (birth cohort about 1900-14)	5	370	357	346	259	113	—	7.2	—
Bengal Hindu villages (Marriages 1945-46)	6	323	288	282	212	100	33	6.2	—
Amish (Cohort about 1900-20)	7	365	462	251	221	82	14	7.0	6.33
Senegal (Sine-Saloun) (births 1963-65)	8	340	306	260	182	92	28	6.1 (c)	6.7
Sweden (Marriages 1841-1900)	9	319	332	279	226	122	12	6.4	—
French Historical Data									
Crulai (Normandy) (Marriages 1674-1742)	10	428	431	359	319	119	10	8.3	5.6 (b)
Tourouvre (Perche) (Marriages 1665-1714)	11	412	425	378	330	164	11	8.6	6.0 (b)
Bilhères-d'Ossau (Béarn) (Marriages 1740-79)	12	414	400	353	319	165	13	8.3	—
Ile-de-France (Marriages 1740-79)	13	527	515	448	368	144	21	10.1	6.1 (b)
Thézels-Saint-Sernin (Quercy) (Marriages 1700-91)	14	385	335	290	242	67	0	6.6	3.7 (b)

II. MACRODEMOGRAPHIC BACKGROUND

Defining a "natural" level of fertility is not as simple a task as might first appear. It even becomes a purely abstract problem when taken in the following sense: "what would the fertility be if this population did not use any method of fertility regulation?"

Such a question forces us to look for points of reference in populations in which both the following conditions obtain:

(1) absence of all methods of birth control (including sterilization and induced abortion).
(2) availability of satisfactory demographic data.

For historical reasons these two conditions are often mutually exclusive and are found together in only three circumstances:

(1) In contemporary populations whose fertility has not yet declined where, in the general absence of vital records, (1) it is possible to make special surveys.
(2) In late nineteenth century populations which, while not having experienced a decline in fertility, already possessed a good civil registration system (Northern European countries, for example).
(3) In historic populations (prior to the nineteenth century) which can be studied indirectly by using parish registers, genealogies etc.

AGE-SPECIFIC FERTILITY RATES. COMPLETED FERTILITY

Table 4.1 presents some of the findings obtained by the methods mentioned above. It was suggested by the approach and some of the data considered by Henry (1961b) in an article on natural fertility. We have added a column giving the mean number of children *per woman surviving to the age of 45.* (Some of the figures are approximate.) It is interesting to compare this last column with the preceding one, which shows the mean number of children *per woman married at the age of 20 and remaining in the union until the age of 45,* obtained by adding the age specific legitimate fertility rates, multiplied by 0.005.

The difference between the two figures is small where the amount of celibacy is small, the age at marriage very low, and mortality low

(1) With some exceptions like anabaptist sects of North America.

(and the risk of widowhood thus reduced). For example, among the Hutterites and the Amish it is between 10 and 15 per cent. If, on the contrary, many women remain single, if the mean age at marriage is clearly well over 20 years, and if the risk of widowhood is high (and remarriage not immediate), the difference between the two figures may reach 50 per cent. Such is the case with historic France, where a woman reaching 45 years of age had an average of 4 to 6 children, while the average completed family size for a woman who married at 20 years and remained in union until 45 years of age often exceeded 8 children.

Taking into account the proportion of women who remain single and those who die before reaching age 45, the *net reproduction rate* in populations where mortality is fairly high is generally reduced to little above 1.

It is important to realize that to compare the fertility of two populations, independently of the effects of mortality and matrimonial customs, one must be able to calculate *age-specific legitimate fertility rates*. "Legitimate fertility" refers here to the Western type of legal union, or any type of union which persists as a stable and effective union throughout the period on which the rate is calculated.

BIRTH INTERVALS

At first glance there seems to be a simple relationship between fertility rates and birth intervals. Indeed it can be shown, given certain assumptions (homogeneous fecundability, a constant non-susceptible period, and negligible intra-uterine mortality) that after a few years of marriage the average birth interval will tend to equal the inverse of the fertility rate.

Actually, although these assumptions are not fulfilled, the relation holds roughly true between 20 and 35 years, but the problems of obtaining and analyzing information on birth intervals further complicate the analysis. The works by Henry (1953, 1961a), Wolfers (1968), and Sheps *et al.* (1967, 1970, 1972) in particular have enumerated these difficulties. At the very least a reading of these articles suggests that the greatest prudence be used in interpreting the "means of birth intervals".

If a statistic such as a fertility rate (by age or length of marriage) is generally unambiguous, the same cannot be said for the "average birth interval". For example, let us consider a group of marriages with completed fertility, i.e., ones in which the wife is at least 45 or 50 years old and in which the union has never been broken by death or divorce.

Table 4.2. *Birth intervals. Role of birth rank and completed family size*

	Interval between birth (n−1) and (n)										
	0-1	*1-2*	*2-3*	*3-4*	*4-5*	*5-6*	*6-7*	*7-8*	*8-9*	*9-10*	*10-11*	*11-12*
Last interval (D_n) (Families of size *n*)	—	—	—	—	39.8	39.9	39.2	37.2	38.1	34.9	30.4	28.9
Next-to-last interval (D_{n-1})	—	—	—	—	32.9	32.6	33.0	30.3	30.6	29.2	29.4	25.0
Difference between last and next-to-last ($D_n - D_{n-1}$)	—	—	—	—	+6.9	+7.3	+6.2	+6.9	+7.5	+5.7	+1.0	+3.9
Mean interval (I_n) (all final sizes)	13.5	22.0	25.2	26.6	28.8	29.3	28.3	29.0	29.0	29.0	26.7	25.6
Difference ($I_n - I_{n-1}$)	—	+8.5	+3.2	+1.4	+2.2	+0.5	−1.0	+0.7	0	0	−2.3	−1.1
Percentage of intervals D_n included in the calculation of I_n	—	—	—	—	13.6	19.5	19.2	31.6	30.1	33.1	45.0	45.5

N.B. The data for this table are based on a sample of non-contracepting families having five or more children.
Source: Leridon (1968).

According to Table 4.1 one might expect that the means of successive birth intervals, birth rank by birth rank, would rise with increasing birth rank since the age-specific fertility rates show a very sharp drop at the end of the reproductive period, and since birth rank and the age of the mother at the time of the birth are highly correlated. Most often, however, just the opposite is observed. Instead of the birth interval lengthening, there is rather a tendency for it to decrease.

The paradox is resolved if the mean birth interval is recomputed for *each completed family size*. A very clear lengthening of the last one or two intervals is then observed, no matter what the size of the family at completed fertility. With this second method of computation the number of observations is equal for all the birth ranks within each series. With the first method, on the other hand, the weighting is not uniform since there is no interval of a higher rank than n in the families whose size is smaller than or equal to n.

Table 4.2 illustrates this point. The third row shows the difference between the last and next-to-last intervals for each completed family size. Column 5-6, for example, gives the difference D_6-D_5 where D_6 is the interval between the fifth and sixth births in families of size 6, and D_5 is the interval between the fourth and fifth births in the same families. The fifth row shows the differences between successive intervals when families of all sizes are considered together. Finally, the last row gives the percentage of "last intervals" which enter into the computation of each of the figures in the preceding row.

Note that the proportion of "long" intervals *rises* with increasing rank order, while the mean interval (I_n) tends to level off and even to decline:

This example reveals two very important effects:

1. *The truncation effect*, whose various manifestations have been analyzed at length in the articles by Henry (1953, 1961a) and by Sheps *et al.* (1967, 1970). To characterize this effect, let us merely say that the process of reproduction cannot be considered as stationary, homogeneous, and of "infinite" duration. Whether the investigation is terminated at a given point in time (as in the case of a sample survey) or at a given age (even if one continues to 45 or 50 years of age, i.e., to the end of the reproductive period), the length of observation is always limited. This introduces various selection biases.

2. *The importance of weighting* resulting from the sampling method or the method of computation. (See Wolfers, 1968, or Sheps and Menken, 1972).

Let us quote in conclusion the last lines of the article by Sheps *et al.* (1970):

"It is clear that despite the arguments in favor of birth interval analysis, the difficulties of interpreting such analyses are great. Considerably more exploration is needed before we can be certain of the usefulness or interpretation of survey data on closed or open birth intervals.

"It is doubtful whether the current emphasis on data for such intervals as an index of changing natality patterns is justified."

III. THE PHYSIOLOGICAL FRAMEWORK OF REPRODUCTION

Turning to microdemographic analysis brings us back to the physiological bases of reproduction. Without becoming involved in an essentially medical discussion (going too far beyond a strictly demographic analysis would be inappropriate), it is important to sketch the framework for subsequent analyses.

Before we discuss the four components of fertility, let us first glance at a few other aspects of reproduction.

THE LIMITS OF THE FECUND PERIOD: MENARCHE AND MENOPAUSE

The fecund period represents only one part of a woman's lifetime. While the generally accepted limits are from 15 to 49 years, in some cases the beginning of the period is pushed back slightly (to as early as 10 years of age) or the end is extended (up to 55 years of age).

The physiological "indices" of the beginning and the end of the reproductive period are the appearance and the cessation of the menses. These two indices are far from perfect, in the first place, because a high proportion of the first menstrual cycles are anovulatory and, secondly, because, for the same reason (and for other reasons), more often than not sterility precedes menopause by several years. In other words, if the "potential" length of the reproductive period is 35 years, or thereabouts, its actual duration is undoubtedly shorter.

To go back to the beginning of the fecund life, the mean age at *the onset of menarche* varies from one population to another. It may also vary for the same population over time, and among social classes. In most Western countries it has declined rapidly in recent decades. For example, in France today the average age at puberty is estimated at around 13 years, while the estimate for the Punjab, India, is 14 to 15 years (Wyon and Gordon, 1971). The maximum level of fecundity, however, is probably not reached for another five years, i.e. not much before the age of 20.

Less is known about the average *age at menopause*, at least for populations whose lack of vital registration raises the difficulty of

estimating age. In populations of the Western type, periods cease naturally at an average age of 48 to 50 years. However, in many cases the arrival of menopause is artificially accelerated, as for example, following a hysterectomy.

Some surveys suggest that menopause may occur earlier in other populations; in India, for example, the beginning is estimated at 44 to 45 years (Dandekar, 1959). While these findings may reflect errors in estimating ages, they suggest that different populations do differ considerably in average age at onset of menopause. In any event, the age of onset of final sterility will be overestimated if it is considered coincident with the beginning of menopause.

THE OVULATORY CYCLE

The monthly ovulatory cycle provides the basic justification for using the microdemographic approach to study fertility. It has been established that the ovum can be fertilized only during a very short time, estimated at between 24 and 48 hours. While this fact, as well as the principal phases of the cycle, is generally accepted, at least four sources of variation complicate the statistical analysis of the cycles:

(1) the average length of the cycle varies from one woman to another (from 10 to 45 days);
(2) for any one woman, the length of successive cycles is often variable;
(3) the position of the day of ovulation within the cycle varies;
(4) it is likely that all these parameters change as the woman grows older.

A "normal" cycle is generally agreed to last between 26 and 30 days, while the post-ovulatory phase (from ovulation to the first day of the following menses) is around 14 days long. Abnormally long cycles, often anovulatory, are particularly frequent in post-puberty and pre-menopausal phases, however, and this may explain the subfecundity observed at these ages.

In addition to the irregularity of cycles, the impact of *anovular cycles* must also be considered. Different authors estimate that between 5 and 15 per cent of all cycles outside the critical zones (e.g. puberty and menopause) are anovular.

We should point out that microdemographic models rarely take into account either the irregularity of cycles or the existence of anovular cycles; the statistical data are not very reliable, and the

complications which would be introduced into the model would outweigh the expected advantages. In practice the length of a cycle is often taken to be a calendar month.

THE LENGTH OF THE GESTATION PERIOD

The duration of pregnancy is first of all a function of its outcome. A frequency distribution of pregnancies which end in miscarriage constructed according to the length of pregnancy would be equivalent to a table of intra-uterine mortality. Hence, we will confine our discussion to live births.

From their own data Hammes and Treolar (1970) estimate that the average period of gestation for a live-born child is 265 days (38 weeks), counting from the estimated day of ovulation, or 279 days (40 weeks) counting from the first day of the last period. Of the pregnancies studied 72 per cent terminated between 36 and 40 weeks after the estimated day of ovulation.

Out of a group of 100 conceptions distributed evenly over a calendar month (month number 0) the following is observed:

2 live births during month number 7
23 live births during month number 8
66 live births during month number 9
9 live births during month number 10.

It would not be far from the truth to say that all pregnancies terminated by a live birth last 9 months. In contrast, the length of the "non-susceptible period" due to pregnancy varies considerably, often widely exceeding 9 months. Since the variation in the length of the non-susceptible period is so great, it is this factor, rather than the variation in the length of the gestation period, which should be taken into account.

THE COMPONENTS OF FERTILITY

According to the most solid research on the mechanisms of reproduction, the following four factors must be considered in the analysis of fertility:

(1) *FECUNDABILITY*. The preferential time unit is the (average) length of an ovulatory cycle, approximately a month. The probability of conceiving, or fecundability, is defined on this basis. Fecundability reflects in part physiological differences, e.g. sperm counts, and in part behavorial differences, e.g., frequency and timing of intercourse.

(2) *INTRA-UTERINE MORTALITY*. A certain proportion of the con-

ceptions which take place do not produce a live birth. This intra-uterine mortality should be taken into account.

(3) *NON-SUSCEPTIBLE PERIOD.* Each pregnancy which is begun puts the woman in a state of "non-fecundability". This period, which varies in length according to the outcome of the pregnancy and the individual case, is prolonged beyond delivery by post-partum amenorrhea. ([2]) From the moment of fertilization until the first ovulatory cycle (or the resumption of sexual intercourse, whichever occurs later), the woman is said to be in a "non-susceptible period". ([3])

(4) *STERILITY.* Fecundity may be interrupted or prematurely terminated by sterility.

The interest of these four concepts lies not only in their solid physiological basis but also in the fact that each can be studied using the techniques of population analysis. Fecundability may be estimated by examining the monthly probabilities of birth at the beginning of marriage; miscarriages can be estimated through special surveys; the study of intervals between births and between marriage and first birth reveals the existence of the non-susceptible period; parity-progression ratios provide information on how sterility evolves with age.

The fact that each of these concepts is not perfectly *discrete* causes some problems. For example, the border line between fecundability and early intra-uterine mortality is difficult to determine, while the existence of anovular cycles makes it hard to pinpoint the end of the non-susceptible period.

Whether these four components are *independent* is even more doubtful, although until now it has been accepted practically without discussion. It is easy to imagine how a woman subject to repeated miscarriages may be exposed to the risk of premature sterility, or how a weak fecundability may be associated with a high early intra-uterine mortality. These points deserve greater attention in the future.

Each of these components will be studied in greater detail in the next section. Since this discussion will prepare the way for a consideration of controlled fertility, two sources of variation, or types of factors, should be kept in mind:

(1) *physiological factors*, governing the "natural" aspect of fertility;

([2]) Post-partum amenorrhea and the absence of ovulation do not necessarily coincide. However, the most recent studies show that on average, the two are of equal length.

([3]) "Non-susceptible period" is the English translation of "temps mort" first used by Louis Henry (translator's note).

(2) *behavioural factors*, associated with the "control" of fertility. The boundary between the two is not always precise.

IV. FECUNDABILITY

We owe the concept of fecundability to Corrado Gini (1924), who defined it in the following way:

> "I call a *woman's fecundability* the probability that the married woman will be impregnated during the month, given the absence of any Malthusian or neo-Malthusian practices aimed at limiting reproduction." (4)

While not explicitly stated, it seems clear that temporary sterile periods should also be calculated.

When only live births are studied and the above is applied to "children who will be live born", we may speak of "*effective fecundability*". While clearly a more ambiguous notion, since a second component, intra-uterine mortality, is also involved, it is one to which we must often resort. The fecundability of a couple practicing contraception will be called *residual fecundability*.

Setting forth this concept, and noting that fecundability is certainly much less than unity, makes reproduction a random process. Even assuming that all the parameters are invariable and of the same value for all women, women's "reproductive histories" (the timing of births and final number of children) may differ greatly. Despite the fact that the first analytic models of fertility were based upon these very restrictive assumptions, they permitted a certain amount of progress to be made.

Given these assumptions, the first conceptions (starting with the first month of exposure to risk) are distributed according to geometric law. The mean time to conception, when computed by counting the number of ovulations before conception (including that ovulation which was followed by fertilization), is

$$M = \frac{1}{p} \qquad (1)$$

The variance of this distribution is:

$$\sigma^2 = (1-p)/p^2 \qquad (2)$$

where the monthly probabilities of conception are all equal to p.

(4) "Malthusian population" refers to a population in which the individual or the couple deliberately attempts to control the number and spacing of their children. The more common English term is "contracepting population" (translator's note).

In reality successive birth probabilities at 9, 10, 11, and subsequent months decrease appreciably. After 8 or 9 months the probability may fall to half its initial value. These observations point to the existence of a process of selection within a heterogeneous group. In fact, natural fecundability in non-contracepting populations and, to an even greater extent, residual fecundability in contracepting populations, is not homogeneous. Thus the problem of how to estimate its distribution arises.

METHODS OF ESTIMATING FECUNDABILITY

Certain problems are involved in the estimation of fecundability. In the first place, it is impossible to make a complete count of conceptions, since the women involved may be unaware of them for some time, or even forever if the pregnancy is terminated by a very early miscarriage. It is thus necessary to proceed retrospectively, starting from the confinements. Most often only live births can be recorded.

The time lag between conception and birth creates another problem. As the distribution of durations of pregnancy given above indicate, it is impossible to allocate the live births observed during a given month to the precise month of conception. Further complications are introduced by premarital conceptions and the position of the day of marriage within the cycle.

All these problems have a far from negligible effect on the estimation of the probability of conception during the first month of exposure which, as we shall see, is the most important.

Gini proposed the following formula for calculating the mean fecundability of women who conceive during the first n months of marriage:

$$\bar{p}_n = 1 - \frac{\sum_{0}^{n-1} C_{x+1}}{\sum_{0}^{n-1} C_x} \qquad (3)$$

This estimate is biased since the fecundability of women who conceive the most quickly exceeds the average for the entire group. But if the summation is pushed far enough, the following is obtained:

$$\bar{p} = 1 - \frac{C_1 + C_2 + \cdots}{C_0 + C_1 + C_2 \cdots} = \frac{C_0}{C} \qquad (4)$$

(setting $C = \sum_{0}^{\infty} C$)

C_0/C is none other than the *first monthly probability of conception*, which, while no longer subject to the selection bias noted above, is,

as we have already mentioned, subject to other biases. The probability of birth of month 9 underestimates by about 15 per cent the actual effective fecundability (Vincent, 1961: 210). This underestimation is often very obvious, as when the probability for month 10 does not differ substantially from the probability for month 9 (Henry, 1965).

The variance (V) of the fecundability distribution can be estimated by the following formula:

$$\bar{p}\left(1+\frac{V}{\bar{p}^2}\right)=1-\frac{C_0}{C_1} \tag{5}$$

Methods of estimation based on mean intervals or fertility rates cannot readily be used since they generally allow the calculation of *a harmonic mean* ([5]) rather than the arithmetic mean of fecundabilities. In addition, they require information on the distribution of the length of non-susceptible periods, which is rarely available.

The same holds true *a fortiori* for methods based on a "combinatorial" analysis of the whole reproductive period. Fecundability is only one of the sources of randomness, and to calculate it with precision, one must know all the others (e.g., non-susceptible period, intra-uterine mortality, sterility).

Yet another method of estimation is by the *frequency distribution of conceptive delays*. If it is accepted, as Henry suggests, that fecundabilities are distributed according to a Beta function (or Pearson's Type I distribution), a bi-univocal relationship can be established between the parameters of this distribution and those of the distribution of conceptive delays.

Let f(p) be the probability density of fecundability:

then

$$f(p)=\frac{p^{a-1}(1-p)^{b-1}}{B(a_1 b)} \tag{6}$$

when:

$$\begin{gathered}0<p\leqslant 1\\ a \text{ and } b>0\end{gathered} \tag{7}$$

$$B(a,b)=\int_0^1 p^{a-1}(1-p)^{b-1}\,dp$$

The mean and the variance of this distribution are

$$\bar{p}=\frac{a}{a+b} \tag{8}$$

$$V=\frac{ab}{(a+b)^2(a+b+1)} \tag{9}$$

([5]) The mean conceptive delay is equal to the inverse of the harmonic mean of the fecundabilities.

a and b are estimated from the mean, m and the variance σ^2 of the observed conceptive delays:

$$\hat{a} = \frac{2\sigma^2}{\sigma^2 - m^2 + m} \tag{10}$$

$$\hat{b} = (m-1)(\hat{a}-1) \tag{11}$$

The advantage of this method is that it takes into consideration all first conceptions, instead of only those occurring in the first one or two months. This method is, however, quite sensitive to variations in m and σ^2. While in theory the "maximum likelihood" method is preferable, the calculations which it requires are much more complex (see Sheps and Menken, 1973: 98).

EMPIRICAL FINDINGS

Gini's method, rarely used by anyone but its author, is based on births registered between the eleventh and twenty-third months of marriage. It consequently excludes conceptions which occur during the first month of marriage or result in births registered after the second year of marriage. Using this method, Gini obtained very similar estimates of effective fecundability for such diverse populations as Berlin, Australia, and Italy. The published values range from 0.18 to 0.21.

Using the probability of birth in the tenth month of marriage, Henripin (1954) obtained a value of 0.25 for Québec, while Vincent (1961) obtained a value of 0.22 for a group of large French families (an estimate which Henry suggests should be raised to 0.25 to take into account the biases mentioned above). For a set of observations taken from historic France, Henry (1965) calculated a probability of 0.20, which should be corrected by the same proportion as was Vincent's estimate.

Estimates using the "method of moments" (Beta distribution) are more recent. Potter and Parker (1964) used a sample of American women who did not practice contraception at the beginning of marriage. They obtained a mean value of fecundability of 0.14 and a standard deviation of 0.076. Jain (1969a) found values of 0.16 and 0.086 respectively in Taiwan. From Vincent's data, we have used the method of moments to obtain the following estimates: 0.23 (instead of 0.22) for the mean, and 0.105 for the standard deviation.

The first two estimates cited in the preceding paragraph are

considered rather low by their authors. (6) The other estimates are in general agreement.

Recognizing that sources of bias generally cause actual fecundability to be underestimated, we may agree that at around 25 years of age effective fecundability is in the neighborhood of 0.25. This is only a provisional order of magnitude, which leaves open the question of variation with age or from one group to another—a topic which we will take up next.

FECUNDABILITY FOLLOWING THE CESSATION OF CONTRACEPTION

Until now the analysis has been based on the distribution of conceptions (or births) following marriage. Would it not be possible to use other distributions?

The first idea that comes to mind is to study the distribution of births after each confinement. However, as we shall see, the existence of a period of post-partum sterility of variable length makes this very hard. Even if the duration of amenorrhea were studied using surveys on individual women, the problem would not be solved, since the period of amenorrhea does not coincide exactly with the period of sterility.

Another idea would be to take advantage of the periods of interruption of contraception which occur in populations where very effective contraception is practiced. It is theoretically quite simple to compose cohorts of women "newly exposed to the risk" of conceiving by using retrospective interviews. In practice, it is difficult to define precisely the first month of exposure, and it is very likely that fecundability will be overestimated.

Some women might be tempted to "rationalize", after they occur, conceptions which were not actually planned. Other conceptions may reflect the perfect planning of women who used their knowlegde of the fecund periods with their cycles in order to become pregnant. In both cases conceptions "during the first month" are likely to be much more numerous than they would in a "neutral" situation, i.e., without either contraception or attempts to conceive. Potter and Parker (1964) in the study cited above compared fecundability "after cessation of contraception" with fecundability "at the beginning of marriage": 40 per cent of the women conceived in the month after contraception was discontinued as against 28 per cent in the first month of marriage. For these various reasons it is difficult to estimate natural fecundability after contraception has been discontinued.

(6) Due to the fact that periods of contraception were not completely eliminated in the American sample and to various biases in the Taiwan study.

VARIATION WITH AGE

Let us return to the analysis of the mean contraceptive delay at the beginning of marriage. A classification based on age at marriage permits us to study the evolution of fecundability in the first half of the fecund lifetime (between 15 and 30 years).

Vincent (1961) observed that fecundability doubles between 15 and 20 years of age, rising only slightly thereafter. Jain (1969a) obtained comparable results by fitting a Beta distribution to each age-at-marriage group. For women married before 16 he obtained a fecundability of 0.09, and for women of 20, 0.18. He found fecundability attained its maximum at 0.22, for women aged 21-25.

While there is little doubt that fecundability increases between the ages of 15 and 20 (or a little later), the situation after the age of 30 or 35 is less clear. It was long accepted that fecundability declined at the end of the fecund period (finally reaching zero), even though the effects of final sterility acquired at these ages could never be completely eliminated. Analysis of fertility rates of women later proven fertile, like that of birth intervals classified according to completed family size, does indeed show that fertility declines before final sterility is attained.

But which of the three components is responsible for this evolution? The information available on intra-uterine mortality and the length of post-partum amenorrhea implicates both of these components, while a study based on a model by Leridon (1973) suggests that the same holds true for fecundability.

Let us also note that the long accepted hypothesis that the fertility of non-contracepting couples is a function of age alone, and not of marriage duration, should be modified today. Research increasingly supports the conclusions that, with age held constant, fertility is higher during the first years of marriage (see Henry, 1972).

DIFFERENTIAL FECUNDABILITY

Excluding the differences due to age, are the differences between the estimates of the average fecundability of various groups significant? Before this question can be answered affirmatively, the effect of various biases must be eliminated:

(1) *THE MEMORY EFFECT.* The longer a woman has been married when she is questioned about her first conception, the greater the chance that she will omit a miscarriage which occurred before the first live birth.

(2) *THE TRUNCATION EFFECT.* The shorter the time a woman has been married at the time of the interview, the more the sample will be biased towards women of high fecundability, since women of high fecundability have a higher chance of conceiving in a given time period (e.g., one year, two years). This effect acts in the opposite direction from the preceding one, but does not necessarily offset it exactly.

(3) *THE IMPORTANCE OF PREMARITAL CONCEPTION.* In principle, they are not included in the analysis. However, it may be difficult to determine whether a confinement after 8 or 9 months of marriage is the result of a post-nuptial or pre-nuptial conception.

(4) *PERIODS OF CONTRACEPTION.* Where they occur, they should also be excluded from the analysis. However, as we have already said, it is difficult to be sure of the accuracy of statements on this topic.

(5) *THE RATE OF OMISSION OF EARLY MISCARRIAGES.* The rate of omission may vary greatly depending on the importance which the respondent attaches to a miscarriage (and also according to her understanding of the mechanisms of reproduction).

After examining a number of these possible biases, Jain (1969b) concluded that in his sample there was a positive correlation between fecundability and socio-economic status (e.g., level of education, husband's profession, and geographic or ethnic origins). Still the author feels that actual physiological differences cannot be confirmed definitely until better information is available on the time required for the consummation of marriage, coital frequency, and the desire to conceive a first child rapidly.

It is very difficult, even impossible, to completely separate *physiological* factors from *behaviorial* ones. The concept of fecundability is relatively simple—and very efficient—as long as we remain at a certain level of analysis. But we have gone beyond the scope of its applicability as soon as we consider that the probability of conception is not the same every day during a month. This is especially true if there is reason to suspect that the population under study is taking this internal heterogeneity into account.

The notion of fecundability has shared the fate of many concepts which, first considered strictly "probabilistic", little by little lose their probabilistic sense, finally becoming simple empirical representations of more complex facts.

V. INTRA-UTERINE MORTALITY

PROBLEMS OF DEFINITION AND OBSERVATION

Consideration of intra-uterine mortality was for a long time limited to the statement that a small percentage of pregnancies do not end in a live birth. The lack of interest in this aspect of human reproduction undoubtedly stemmed from the difficulties of obtaining accurate information and from the widespread confusion (even in medical circles), between spontaneous and induced abortions. We will come back to the problem of obtaining data when we discuss the computation of various "tables". First, let us turn briefly to the question of induced abortion.

Induced abortion, whether legal or illegal, is statistically important in many countries. Even in countries where abortion has been legalized with certain restrictions, illicit abortions may continue to be performed in a non-negligible number. Abortion cannot be considered a "contraceptive method", because it terminates an already existing pregnancy. This gives it its specificity and, it may even be said, its strength.

When induced abortion is illegal, and the declaration of spontaneous abortions not compulsory (which is generally the case), data on intra-uterine mortality can be obtained only by survey, usually retrospectively. The usual problems associated with this method of inquiry are compounded by the difficulty of interpreting respondents' answers. Distinguishing between spontaneous abortions and disguised induced abortions is often difficult. Several different situations may arise:

(1) The respondents have no scruples about revealing their induced abortions. This is the best possible situation, but one can rarely be certain that it holds true for all the women surveyed. In Hungary, for instance, women interviewed in 1966 reported only 60 per cent of the actual number of legal abortions performed during the same period of time.

(2) The respondents avoid mentioning their induced abortions, declaring only spontaneous abortions. While confusion is thus avoided, the various estimates (e.g., conceptive delays) will nonetheless be biased.

(3) The respondents disguise their induced abortions as miscarriages. Clearly this will distort the estimate of spontaneous intra-uterine mortality. In instances where disguised induced abortion is significantly high, the estimate of spontaneous abortions will be so

patently inflated that one should be able to infer an estimate of disguised reports.

A novel method of estimation has recently been proposed by Warner (1965) and used by Abernathy *et al.* (1970) and by Chi *et al.* (1972). Each woman interviewed is given an urn containing a known proportion of white and black balls. She is asked to draw out one ball and, without showing it to the interviewer, to answer one of two questions, depending on the color of the ball drawn:

(1) For a white ball: "Were you born in the month of January?"

(2) For a black ball: "Did you have an abortion within the last twelve months?"

The interviewer records the answer (yes or no) but does not know which question has been answered.

Let:

P equal the proportion of "yes" answers.

N_1 equal the number of women drawing a white ball.

N_2 equal the number of women drawing a black ball.

P_1 equal the proportion of women answering "yes" to the first question.

P_2 equal the proportion of women answering "yes" to the second question.

$$N = N_1 + N_2$$

Then

$$P = (N_1 P_1 + N_2 P_2)/N \qquad (12)$$

The mathematical expectation for P,

$$E(P) = E(N_1/N) \times E(P_1) + E(N_2/N) \times E(P_2) \qquad (13)$$

(if the various probabilities are considered independent).

Now,

$E(P_1)$ can be estimated from vital statistics or census data,

$E(N_1/N)$ and $E(N_2/N)$ are known (from the relative proportion of white and black balls in the urn),

$E(P)$ can be estimated from the proportion, P^*, actually obtained.

From this, one can deduce the value of $E(P_2)$.

While this method may completely satisfy the statistician, it is uncertain whether it will inspire the full confidence of the person being interviewed, who after all has rarely been schooled in probabilistic reasoning! Nevertheless, it is a method which deserves to be more widely tested, especially with truly representative samples.

Let us end this section by clarifying a few terms:

(1) The "period of gestation" may be measured from the first day

of the last menstrual period (this is the "customary period" commonly used in English-speaking countries) or from the presumed day of conception (this is the "true period" or "true length" of gestation). There is about two weeks' difference between the two measures.

(2) Although we will speak of the "intra-uterine mortality table", the event studied is always the expulsion of the conceptus (the termination of the pregnancy), and not the death of the embryo, which may occur much earlier.

(3) In countries with a vital registration system the declaration of miscarriages is usually compulsory for still-births only, usually defined as fetal deaths over the conventional length of gestation of 28 weeks (although in some countries the definition also includes fetal deaths from 20 to 28 weeks).

COMPUTATION USING A LIFE-TABLE ANALYSIS

The general intra-uterine mortality rate which can be obtained through simple retrospective interviews is around 10 to 15 per cent. Similar figures can be obtained from prospective methods of investigation if one is satisfied with relating the number of miscarriages to that of registered pregnancies. Both methods seriously underestimate intra-uterine mortality.

In contrast to the classic tables, such as life tables, the points of entry into observation are in this basis spread over time, since on the first day of her pregnancy no woman can be sure that she is pregnant. Thus, it is impossible to construct a "cohort" in the demographic sense of the word, i.e., "a group of persons who experience a certain event in a specified period of time." ([7]) However, one can construct a table by the method of probabilities, as long as care is taken to record *the gestational age at the time the pregnancy is first diagnosed.*

The idea was first applied by Drs. French and Bierman (1962). Using the following notation:

D_t equals the number of miscarriages in the gestational interval ([8]) t to t+1

N_t equals the number of live births in the gestational interval t to t+1

([7]) The time scale is not necessarily tied to a calendar year. It may be a time "natural" to the phenomenon under study, e.g., the duration of marriage. Contrary to this definition, taken from the *Multilingual Demographic Dictionary* (New York, United Nations, 1958, item 116.2) it is not even always necessary to refer to a period of time. The cohort may be defined simply by the occurrence of the event, as long as it is not ambiguous (an example is the group of women having given birth to their first child).

S_t equals the withdrawals from study by departure or death of the woman between t and t+1

E_t equals the number of women first reported pregnant (i.e., pregnancies diagnosed) in the interval t to t+1.

The number of pregnancies under observation at time t is

$$G_t = G_{t-1} + E_{t-1} - (D_{t-1} + N_{t-1} + S_{t-1}) \tag{14}$$

The estimated probability of intra-uterine mortality in the interval t to t+1 (which is in fact a probability of expulsion) may be computed in either of two ways, by treating a live birth as (a) a *disturbing* event (thus calculating the probability of occurrence of a single event) or (b) a *competing* event (thus calculating the probability of occurrence of two dependent events). The second approach is the more logical, since the two events are truly competitive, i.e., mutually exclusive: the occurrence of either one precludes the other. (9)

Thus, it is really necessary to compute two probabilities, one for each possible outcome:

(1) the probability of aborting

$$q_t = \frac{D_t}{G_t - 0.5\,(S_t - E_t)} \tag{15}$$

(2) the probability of delivering a live-born child

$$V_t = \frac{N_t}{G_t - 0.5\,(S_t - E_t)} \tag{16}$$

There are likewise two categories of "life table events": deaths (d_t) and live births (n_t), and the "survivors" (l_t) who escape both events:

$$d_t = q_t\, l_t \tag{17}$$

$$n_t = V_t\, l_t \tag{18}$$

and

$$l_{t+1} = l_t - d_t - n_t \tag{19}$$

The formulae for q_t and V_t are based on the following assumptions:

(8) The "gestational interval" is the period of gestation between two exact gestational ages stated in weeks.

(9) The first solution can be justified by the fact that it is less affected by the choice of time-unit, (month, week or day); thus the result is closer to the concept of "instantaneous rate", particularly at the beginning and at the end of the table.

Table 4.3. *Intra-uterine mortality by gestional age*
[*Life table* (a) *probabilities and deaths per 1,000 pregnancies*]

Conventional gestational age (in weeks)	French and Bierman (15) (c)		Taylor (16) (c)		Petterson (17) (c)		Shapiro et al. (18) (c)		Erhart (b) (19) (c)	
	q_t	d_t	q_t	d_t	q_t	d_t	q_t	d_t	q_t	d_t
0									0.112	112
4	0.108	108	0.061	61	(0.016) (d)	(16)	0.014	14	0.082	73
8	0.070	62	0.049	46	0.064	63	0.059	58	0.067	55
12	0.045	37	0.025	23	0.044	40	0.040	37	0.028	21
16	0.013	10	0.011	10	0.006	5	0.014	12	0.011	8
20	0.008	6	0.008	7	0.001	1	0.006	6	0.009	7
24	0.003	2	0.003	3		0	0.004	3	0.002	2
28	0.003	2	0.004	3			0.002	2	0.004	3
32	0.003	2	0.003	3			0.003	3	0.002	1
36	0.004	3	0.004	3			0.004	3	0.007	5
40	0.007	5	0.004	4			0.005	2	0.011	8
Sum of deaths (weeks 4-39)		232		159		125		138		175
Number of deaths observed (weeks 4-39)		262		729		116		872		—

(a) These are "double decrement" tables, where exit is either by live birth or fetal death. Only the probabilities and deaths which correspond to the latter are included.

(b) These figures have been adjusted and extrapolated by the author according to his own data.

(c) Numbers refer to references.

(d) Week 7 only.

Sources: See end of chapter.

(1) The different events considered are statistically independent.

(2) If each type of event were considered alone, it would be evenly distributed within each interval $(t, t+1)$. Moreover, the probabilities which are computed are low.

The main problem arises in the beginning of the table. During the first weeks, the "entries" into observation are sometimes far from being evenly distributed over time. Under these circumstances, the correction factor,

$$0.5\,E_t$$

is inaccurate. For example, when separate computations are made by week and by month (of 4 weeks), the difference between the two monthly probabilities obtained may be as high as 100 per cent for the second month of gestation! Thus, it is advisable to compute at least the beginning of the table on a weekly basis.

The table will never begin at $t = 0$. In practice, nothing can be expected before 4 weeks of the conventional gestational period has elapsed, i.e., at the time when the first menstrual period is missed. The quality of the table which is obtained will depend heavily on the number of entries registered in the intervals between the fourth and eighth week of the gestational period. In this respect, the table of French and Bierman remains one of the very best.

EMPIRICAL FINDINGS: PUBLISHED TABLES

Table 4.3 gives several sets of estimates of probabilities and deaths, some of which we have reconstituted from data published in a different form. The deaths represent the number which would be recorded per 1,000 (known) pregnancies in progress at 4 weeks, measured from the first day of the last period.

The sum of the deaths in each column, shown on the next-to-last line, varies from 125 to 232 per 1,000. It should be noted that most of the difference between the upper and lower figures can be explained by deaths occurring in the first four weeks (4 through 7) covered by the table.

The estimates vary within a much smaller range when the summation of deaths is started with week 8 instead of week 4. The totals obtained in this way vary between 98 and 124 per 1,000.

For this reason French and Bierman's estimate may be considered the best: *out of 1,000 pregnancies in progress at the beginning of the third week* (of the "true" gestation period), *200 to 250 will be terminated by a spontaneous abortion* (or a still birth).

This conclusion obviously assumes that the data on which the

tables are based include only spontaneous abortions. Theoretically, all the authors were concerned with this problem. In addition, had a high proportion of induced abortions been included, there would be an anomaly around weeks 8 to 11, the time at which the majority of induced abortions are performed.

THE ROLE OF THE MOTHER'S AGE, THE PREGNANCY RANK, AND OUTCOME OF PREVIOUS PREGNANCIES

As we have already mentioned, the mean intra-uterine mortality rate obtained by direct observation (i.e., not calculated from the table) is usually between 10 and 15 per cent. This average obtains when all ages and ranks are taken together. If pregnancies are classified according to the age of the mother, a sharp rise in the rate is always observed beginning at 30-35 years. As a rule the rate doubles between 20 and 40 years.

The situation is less clear for the very early ages (before 20 years). According to some studies, the rate at 15-19 years is higher than at 20-24 years. But other studies suggest that the rise is continuous after 15 years of age.

Supported by some ten studies, we propose the following average series, based on an average rate of 150 per 1,000 as resulting from these retrospective studies.

Age-group	*Less than 20 years*	*20-24 years*	*25-29 years*	*30-34 years*	*35-39 years*	*40 years and more*	*All ages*
Rate per 1,000	118	122	135	160	200	270	150

If we accept the true mean rate to be the one derived by the life-table analysis of French and Bierman (237 per thousand), the series becomes:

Age-group	*Less than 20 years*	*20-24 years*	*25-29 years*	*30-34 years*	*35-39 years*	*40 years and more*	*All ages*
Rate per 1,000	186	193	214	253	316	426	237

Data classified both by age of mother and pregnancy rank confirm this tendency to vary with age. The role played by pregnancy rank is quite complex. Indeed, it seems that one must consider not only

the rank order itself, but the entire reproductive history. The most striking fact is that *the risk of a miscarriage doubles as soon as a woman has had at least one*, and especially immediately after this first miscarriage. In a systematic study of a sample of 4,400 pregnancies, ranks 1 to 6, (Leridon, 1973: 65-74) we concluded that:

(1) The risk is double, no matter what the rank of the first miscarriage;
(2) The rate then remains at a very high level for at least three or four successive ranks; we are not dealing with a "cluster effect" limited to two or three pregnancies;
(3) Nor can it simply be the result of pre-existing heterogeneity (although this heterogeneity certainly exists);
(4) The effect of age should not be confused with this particular effect.

Several explanations for this phenomenon may be considered. It is possible that whatever caused the first miscarriage (e.g., disease or local trauma) may become definitively established without other visible effects which would justify a visit to a doctor or other medical treatment. It is also possible that the first miscarriage itself increases the risk (through the after-effects of a bad curetage or the absence of any treatment). We can only venture hypotheses in an area which as yet has been little studied.

THE PROBLEM OF DIFFERENTIAL INTRA-UTERINE MORTALITY

The numerous difficulties involved in obtaining accurate information make it very difficult to compare results between different social groups or different countries. All other things being equal, it is clear that the better the investigation, i.e., the greater the number of early miscarriages that are detected, the higher the observed rate will be. The number of miscarriages which are declared depends on a series of interrelated factors, especially psychological ones, which reflect the importance attached to this type of event: the suspicions which their declaration may well arouse in the interviewer ("Aren't these really attempts at concealing induced abortions?"; "Isn't this really a sign of 'poor health'?"); the taboos which surround certain aspects of reproduction... Not to mention, of course, the objective difficulties involved in recording (for example, confusion between an amenorrheic cycle and a miscarriage).

Among the difficulties encountered in an Algerian survey, F. and M. von Allmen-Joray (1971) cite the existence of the myth of the "sleeping child": when a woman who knows (or believes) she is

pregnant does not carry her pregnancy to term, it is because the child "fell asleep" in her womb; he will "wake up" at the time of the following pregnancy, 6 months or a year later. In such circumstances the very notion of intra-uterine mortality is denied! Without going to such an extreme, it must be admitted that in these occurrences there is a certain subjectivization of a scientifically undeniable phenomenon.

That being the case, a differential intra-uterine mortality may still plainly exist. For example, Retel-Laurentin (1973) has noted that the N'Zakara (Central African Republic) and the Bobo (Upper Volta) have an extremely high incidence of abortions (and sterility), a situation which reflects a particularly unfavorable collective pathology.

A considerable amount of data is available on perinatal mortality in the developed countries. In these studies, only pregnancies lasting at least six months are considered. While there is good reason to believe that the etiology of early miscarriages differs quite sharply from that of late miscarriages, the aspects of the etiology which may be similar commend these studies to our attention.

These studies of still-births generally confirm the findings which we have cited for intra-uterine mortality concerning the role of age, rank, and previous obstetrical history. They reveal further differences according to social class, and linked with this, the intensity of prenatal care (e.g., number of medical examinations, date of first examination). Thus, it seems reasonable to believe that any policy of perinatal health also affects late intra-uterine mortality, though to a much lesser degree.

INTRA-UTERINE MORTALITY AND FECUNDABILITY

As we have seen, the intra-uterine mortality tables which it has been possible to construct, do not take into account expulsions occurring during the first two weeks after conception. There is every reason to believe, especially after the studies of Hertig (1967), that a considerable number of such expulsions occur. "Fecundability" as defined by demographers on a monthly basis thus incorporates a certain amount of intra-uterine mortality. This does not pose a serious problem as long as the risk of fertilization is not modified by the existence of a very short-lived pregnancy. If this is really the case, it amounts to saying that there was no fertilization, or that the fertilization took place but without having any effect on the following cycle.

This contention cannot be supported by direct observation. How-

ever, we know that for miscarriages occurring after two or three months of gestation, the non-susceptible period is on the average not more than one or two months longer than the gestation period. It is likely that the same holds true *a fortiori* for shorter periods and that the classic notion of fecundability remains very useful.

Let us also point out that a recent article (Barrett and Marshall, 1969) has given the first estimate of daily probabilities within the cycle. These probabilities are significantly different from zero for the five days preceding ovulation and the first day after ovulation.

When combined with a frequency distribution of coitus (by month), they produce a monthly probability of conceiving which is entirely consistent with the usual estimates of fecundability.

VI. THE PHYSIOLOGICAL BASIS OF THE NON-SUSCEPTIBLE PERIOD

Any conception marks the beginning of a period in which fecundability is nil, a period which lasts at least as long as the gestation period (and often much longer), and which is called a non-susceptible period.

POST-PARTUM AMENORRHEA AND BREAST-FEEDING

The existence of a period of post-partum sterility linked to the practice of breast-feeding is a fact of everyday observation. In all societies, even those where the least is known about the mechanisms of reproduction, anyone can see that the risk of another conception is much higher when the previous child dies very young than when he survives and is breastfed. Some even think that breast-feeding is sometimes prolonged solely because of its contraceptive function.

For a long time, however, reliable statistical data on this subject were lacking. The first attempts at empirical verification were made only indirectly. But let us turn first to the direct observation of the duration of amenorrhea.

In the complete absence of breast-feeding, the average period of post-partum amenorrhea lasts around two months. It increases as the period of breast-feeding lengthens, although at a decelerating rate.

for 3 months of breast-feeding: around 3 months of amenorrhea
for 6 months of breast-feeding: 4 to 5 months of amenorrhea
for 9 months of breast-feeding: around 6 months of amenorrhea.

When breast-feeding is extremely prolonged (two years or more),

the duration of amenorrhea appears to reach a maximum somewhere around 15 months. Under the most common circumstances, where the average duration of breast-feeding is about 18 to 24 months, the average duration of amenorrhea is in the neighborhood of 12 months.

Is this "duration of amenorrhea" exactly synonymous with "duration of post-partum sterility"? In theory, no. For a long time it was generally agreed that the period of amenorrhea might be lengthened by one or more cycles which were anovular and hence sterile (Tietze, 1963). The most recent data throw these assumptions into question: in about half the cases, the first ovulation takes place *before* the return of the menses, and hence during the period of amenorrhea (Pascal, 1969; Perez *et al.*, 1971). The result may depend on duration of amenorrhea, but on the average, we can agree, the duration of sterility is very close to the duration of amenorrhea.

The return to normal fecundity is probably not sudden: the first cycles are often irregular (high variance) and some may even be anovular.

INDIRECT MEASURES: BIRTH INTERVALS

An indirect estimate of the duration of post-partum sterility may be obtained by comparing the interval between marriage and first birth with intervals between births. (Data must be for non-contracepting populations.) The two types of intervals have in common:

(1) the period of gestation (which does not vary greatly);
(2) the conceptive delay (which varies little from one birth order to the next).

They differ from one another in the existence of post-partum sterility.

The difference between the "marriage-first birth" interval and the "first-second birth" interval varies between 5 and 12 months—a finding which agrees with what we have seen concerning the length of breast-feeding and amenorrhea.

An idea of how much breast-feeding lengthens the non-susceptible period can be obtained by classifying the birth intervals according to age of the first of two children, or according to the age at weaning (although when breast-feeding is generally practiced, voluntary weaning is rare before six months or a year).

The immediate result of the child's death is in fact to stop the breast-feeding and to end the inhibition of ovulation connected with breast-feeding. Many studies offer such comparisons, between intervals "after death" (of a child dying before 3 months, or 6 months, or a year), and for survivors.

Table 4.4. *Birth intervals according to longevity of child born at beginning of the interval*

	Mean interval (in months)						
	Québec (2) (a)	*Crulai* (10) (a)	*Tourouvre* (11) (a)	*Senegal* (20) (a)	*Mömmlingen* (21) (a)	*Taiwan* (22) (a)	*Cocos* (23) (a)
Child dying before 3 months	18.8	20.7	20.7	19.7	16.1 (b)	11.4 (d) +9	20.4 (b)
Child dying between 3 and 6 months				21.5	17.0 (c)		20.0 (c)
Child dying between 6 and 9 months	23.5			22.1	20.2		25.7
Child dying between 9 and 12 months				25.1			
Child surviving to a year	25.0	29.6	30.0	32.7	27.5	17.3 (d) +9	29.3

(a) Numbers refer to references.
(b) Deaths between 0 and 1.2 months.
(c) Deaths between 1.2 and 6 months.
(d) Interval calculated between delivery and the following conception (add 9 months to compare this with the other observations).
Sources: See end of chapter.

The results cannot always be compared directly with each other, for the reasons which we mentioned above (page 98):

(1) The length of observation time is often limited: this also applies to the analysis of the frequency distributions of the length of breast-feeding and amenorrhea, which are sometimes highly skewed toward the left and must be truncated. (This explains, for instance, the use of medians instead of means, and the method of computing "survivorship functions".)

(2) The means of the intervals can be calculated in very different ways, e.g., means by rank, by family, weighted or not weighted... To cite only one example, it is clear that the mean is very sensitive to the inclusion or exclusion of the last interval of each family which, as we have seen, is definitely longer on the average than the preceding ones.

For these various reasons, the findings in Table 4.4 are not completely comparable. But this point is secondary. More important is the fact that they all confirm the expected tendency: depending on whether the child died before 3 months, or survived to beyond a year, the mean interval differs by between 6 and 13 months.

EVOLUTION WITH AGE AND THE STUDY OF VARIANCE

Taking into account the age of the mother raises a new problem of selection. A birth interval is not a momentary event for which age may be defined unambiguously. Some convention must be adopted for defining the age of the mother. If we choose the age at the beginning of the interval, we face a severe bias at the end of the fecund period: only short intervals will be able to be represented at the higher age levels. If, on the other hand, we chose the age at the end of the interval, a similar bias will obtain at the beginning of the fecund period. To overcome these two biases, one must either keep a sufficient distance from both the upper and lower limits, or use the age at the beginning of the interval for the beginning of the fecund period and the age at the end of the interval for the end of the fecund period.

From a survey conducted in the Punjab, India, Potter *et al.* (1965) concluded that the interval increases by 20 per cent between 20-24 and 35-39 years of age. Their results closely match those obtained earlier by Henry (1958), using both direct and indirect methods (based on the inverse of the fertility rates of women later proven fertile).

In their study, Potter *et al.* explain a little more than *half this*

increase by the lengthening of the period of post-partum amenorrhea with age.

The analysis of variance can be carried out in two different directions: either to bring out the role which each of the components (non-susceptible period, conceptive delay, pregnancy wastage) plays (the approach of Potter *et al.*), or to study intra- and inter-family variations (cf. Henry, 1958).

The article by Henry showed that, based on a limited sample:

(1) The greatest part of the total variance results from inter-familial differences.
(2) The greatest part of the *intra*-familial variance can be attributed to the random nature of the process of reproduction.

In other words, there are "long interval" families (with low fecundity) and "short interval" families (with high fecundity). For a given mean interval, however, the variances differ only because of the random nature of the process of reproduction.

The importance of this "random variance" is also brought out in the article by Potter *et al.*; they conclude that the variance of the length of amenorrhea explains only a quarter of the total variance of the intervals.

VII. FINAL STERILITY

INTRODUCTION: STERILITY AND INFERTILITY, DEFINITIONS

If the birth of a child indisputably proves that a woman and her spouse are fecund, the inverse situation is open to several interpretations. Between "certain fecundity" and "certain sterility" there is a whole range of intermediate conditions, which can be explained as follows:

(1) When a couple does not succeed in having a child, the physiological responsibility may lie with the wife, the husband or both. In the extreme case, it is possible that each of them individually may be fecund, yet the couple be sterile. Doctors now tend to divide the responsibility for pathological sterility detected at the beginning of marriage more or less evenly between the sexes.

(2) Sterility may be expressed not as an inability to conceive, but as an inability to successfully bring a pregnancy to term (due to repeated miscarriages). Although their etiology is undoubtedly different, they cannot be distinguished on a statistical basis, since only live births (and, under the most rigorous conditions, stillbirths) are registered in most countries.

(3) Even with a fecund couple, conception does not necessarily take place during the first cycle of exposure to risk. Quite a long time may elapse if the woman's fecundability is low. As a result, a woman may be infertile after age x while she becomes sterile only after age $x+y$.

Given the great likelihood that there are women with very low fecundability (and this is undoubtedly true of all women at the end of their fecund period), we see that the fecund state and the sterile state are on a continuum.

(4) Similarly, *primary* sterility can be established only as existing at the time of marriage, without knowing whether the woman was previously fecund.

By way of contrast, we will call "secondary" that sterility which is acquired after the birth of one or more children.

For all the above reasons, it is rather arbitrary to draw a sharp line between sterility and fecundity. In surveys of modern populations, where the practice of contraception compounds the problems of analysis (especially when the practice is unequally spread), attempts have been made to define a "fecundity gradient".

Despite the existence of this continuum between sterility and fecundity, it should be stressed that absolute sterility does exist, and indeed that all women become sterile at one age or another.

MEASURING STERILITY IN THE ABSENCE OF CONTRACEPTION

Where contraception is not practiced, a link between sterility and infertility does exist. Let us consider a group of women of age x who have just delivered. These women were all still fecund at age $x-9$ months. A certain proportion (S_x) of this group will never have another child, having become sterile before having had the time to conceive again. If d is the mean conceptive delay after a delivery (delay which here includes the post-partum non-susceptible period as well as the effects of possible miscarriages), then S_x may be considered an estimate of the proportion of women who became sterile between:

$$(x-9 \text{ months}) \text{ and } x+d.$$

This interval includes a gestation and a delivery. S_x thus reflects two different types of sterility risks:

(1) The risk associated with pregnancy, labor and the puerperium (or lying-in period), and

(2) The risk which results from growing older ("spontaneous" or physiological sterility).

Let us suppose that we can classify newly-weds according to age at marriage and can calculate the proportions $S_0(x)$ of couples married at age x (generally using the wife's age) who never have any child. $S_0(x)$ *gives us an estimate of the primary sterility rate at (approximately) age* $x+d$. The complement of S_0 is none other than the "parity progression ratio" of the order 0, which we will write as $a_0(x)$. The sequence of the $S_0(x)$ or the $a_0(x)$ allows us to follow the evolution of the primary sterility rate with age.

This estimate is valid only for married couples. There is reason to believe that the sterility rate at a given age is not the same for married as for single persons. For example, the fact that some marriages are provoked by pregnancy tends to raise the fecundity of married couples. On the other hand, as we have already pointed out, in addition to the effects of age, maternity unquestionably has its own risks, which the bride has not yet experienced.

The parity progression ratio method may also be applied to successive birth orders: $a_n(x)$ represents the proportion of couples ultimately having at least $n+1$ children, among those who had their nth at age x.

It should be noted that in this case the complement of a_n measures the additional cases of sterility acquired between x and the mean age at the $(n+1)$th birth. It reflects the risk due to maternity, additive to the effect of age, which is not included in $1-a_0$.

The available information suggests that the additional risk does not depend on birth rank.

This method measures the incidence of secondary sterility. If we are not primarily interested in making a distinction between primary and secondary sterility, the best method of estimating the total proportion of sterile women at a given age is to compare the (legitimate) fertility rate at a given age with the fertility rate of women who are still fertile after that age. Using five-year age groups, the following rough relationship obtains:

$$\frac{F(x, x+5)}{F'(x, x+5)} = 1 - S(x+2.5) \qquad (20)$$

where F is the fertility rate of the aggregate of women, and F' is the fertility rate for women who are still fertile after age $(x+5)$.

Applying this formula to various populations of European origin (seventeenth and eighteenth centuries), Henry (1961b) obtained the following set of figures:

Age-specific percentages of sterility in married women

	Wife's age (in years)				
	20	25	30	35	40
Per cent sterile	3	6	10	16	31

This series is clearly not universally applicable. Before applying these results to another population, it would be especially advisable to compare the (legitimate) fertility rates observed for ages 35-39, 40-44, and 45-49 with those which appear in Table 4.1.

REMAINING PROBLEMS. THE CASE OF CONTRACEPTING POPULATIONS

None of the methods proposed above (with the possible exception of the proportion of couples who remain childless) enables us to use current statistics to estimate sterility according to age. This is clearly a source of considerable difficulty, and it explains why those very rare data which are available on historic populations are often used even when today's medical, social and cultural context has little in common with that of eighteenth century Europe.

The problem becomes practically insoluble for contracepting populations, since infertility at a given age is only very distantly related to sterility at that age. Under these conditions, surveys, even socio-medical surveys, should be used exclusively. A "fecundity gradient" can then be defined based on information obtained on the desire to have a child, however expressed, and the medical indices of sterility or sub-fecundity. Unfortunately, the results obtained rarely permit us to propose any but a fairly wide "confidence interval" for sterility. On the other hand, this type of survey furnishes interesting information on the importance of fecundity disorders and the efforts couples make to overcome them.

We should also point out that in certain populations or subpopulations there clearly exists a collective pathology of sterility, including chronic gynecological lesions, gonnorrhea and other venereal diseases. This pathology probably also results in an abnormal incidence of intra-uterine mortality.

VIII. INTRODUCING CONTRACEPTION INTO THE MODEL

THE CONCEPT OF CONTRACEPTIVE EFFECTIVENESS

In order to give the concept of contraceptive effectiveness a solid theoretical basis, it is best to define it with reference to fecundability. We consider that "effective contraception" occurs only during those periods when there is actual risk of conception. These conditions obtain for women who are still fecund, outside of their non-susceptible periods and periods of separation.

Given a woman of "natural" fecundability equal to P_N, using a contraceptive method so as to have an actual risk of conceiving each month equal to P_R (residual fecundability), the *effectiveness* E of the contraception is equal to

$$E = 1 - \frac{P_R}{P_N} \qquad (21)$$

The obvious disadvantage of this definition proposed by Tietze and Potter is that it is purely theoretical. As a rule, a woman's "natural" fecundability is not known. For this reason Henry (1968) has related the concept to a classic measure of effectiveness, Pearl's index, standardized for 12 months.

In a cohort of women, newly-weds for example, the following ratio is calculated:

$$R_{12} = \frac{\text{total number of pregnancies registered}}{\text{number of years of exposure to risk}} \times 100$$

which is expressed in "pregnancies per 100 woman-years".

In computing the numerator, all pregnancies are counted, whether or not they were planned. Periods of separation, of pregnancy, and of amenorrhea are not included when computing the denominator. In order to make the index comparable from one group to another, the period of observation is ended at 12 months.

The correspondence between E and R_{12} is the following:

R_{12}	E(%)
1.5 to 2	99.5
3 to 4	99
6 to 8	98
12 to 21	95
around 30	90

WHAT IS "NATURAL FERTILITY"?

In the introduction we defined a "natural system of fertility" as one in which "neither of the individuals, nor the couple jointly, consciously attempts to regulate the number or spacing of their children". According to this definition, behavior which is imposed on the couple and is not a reflection of individual choice does not alter the "natural" quality of a system of fertility. This is often the case with matrimonial customs (e.g., age at marriage, conditions for the remarriage of widowed persons), breast-feeding (even though it is essential to the child's survival), and sexual taboos.

In this sense there is certainly not *one* natural fertility, but many systems of natural fertility, which reflect the influence of socio-cultural factors. Quite considerable differences in fertility can be found even in such apparently homogeneous populations as those of eighteenth century Europe.

Thus, it is impossible to answer the question, "What would be the natural fertility of a particular population if contraception were stopped completely?" We hope that by emphasizing how greatly the components of fertility vary in a non-contracepting population, we will make the reader more cautions in interpreting variations in fertility as being the result of contraception.

REFERENCES

Abernathy, Greenberg and Horvitz

1970 "Estimates of induced abortions in urban North Carolina", Demography 7, 1 (February): 19-30.

Allmen-Joray, F. and M. von

1971 "Condition sociale et fécondité", A.A.R.D.E.S., Secrétariat au Plan, Alger, Algérie: 97-98.

Barrett, J. C. and J. Marshall

1969 "The risk of conception on different days of the menstrual cycle", Population Studie 23, 3 (November): 455-461.

Chi, I. C., L. P. Chow and R. V. Rider

1972 "The randomized response technique as used in the Taiwan outcome of pregnancy study", Studies in Family Planning 3, 11 (November).

Dandekar, K.

1959 "Demographic survey of six rural communities", Gokhale Institute of Politics and Economics.

French, F. E. and J. E. Bierman

1962 "Probabilities of fetal mortality", Public Health Reports, 77, No. 10 (October): 835-847.

Gini, C.

1924 "Premières recherches sur la fécondabilité de la femme", International Mathematical Congress, Toronto: 889-892.

Hammes, L. M. and A. E. Treolar

1970 "Gestational interval from vital records", American Journal of Public Health, 60 (August): 1496-1505.

Henripin, J.

1954 "La population canadienne au début du XVIII[e] siècle," P.U.F., Paris.

Henry, L.

1953* "Fondements théoriques des mesures de la fécondité naturelle", Revue de l'Institut International de Statistique, 3: 135-151.

1958 "Intervals between confinements in the absence of birth control", Eugenics Quarterly 5, 4 (December): 200-211.

1961a* "Fécondité et familles - modèles mathématiques (II)", Population 16 (1): 27-48.

1961b "Some data on natural fertility", Eugenics Quarterly, 8 (June): 81-92.

1965* "Mortalité intra-utérine et fécondabilité", Population 19(5): 899-940.

1968* "Essai de calcul de l'efficacité de la contraception", Population 23(2) (mars-avril): 265-278.

1972 "La fécondité des mariages dans le Sud-Ouest de la France, 1720-1829", Annales E.S.C., 27, 4-5 (Juillet-Octobre): 977-1023.

Hertig, A. T.

1967 "The overall problem in man" in Comparative Aspects of Reproductive Failure, Benirschke (ed.), Springer-Verlag, New York.

Jain, A. K.

1969a "Fecundability and its relation to age in a sample of Taiwanese women", Population Studies, 23, 1 (March): 69-85.

1969b "Socio-economic correlates of fecundability in a sample of Taiwanese women", Demography, 6, 1 (February): 75-90.

Leridon, H.

1968 "Les intervalles entre naissances: nouvelles données d'observation", Population, 22 (September-October): 821-840.

1973 "Aspects biométriques de la fécondité humaine", I.N.E.D.-P.U.F. (Cahiers No. 65), Paris.

Pascal, J.

1969 "Quelques aspects de la physiologie du post-partum", Thèse de doctorat en médecine, Nancy.
(see a note by H. L. in Population 26 (janvier-février 1972, 117-120).

Perez, A., P. Vela, R. G. Potter and G. Masnick

1971 "Timing and sequence of resuming ovulation and menstruation after childbirth", Population Studies, 25, 3 (November).

* Now available in English in: On the Measurement of Human Fertility, Selected Writings of Louis Henry, Elsevier Publishing Company, Amsterdam and New York, 1972.

Potter, R. G. and M. P. Parker
1964 "Predicting the time required to conceive", Population Studies, 18 (July): 99-116.

Potter, R. G., J. B. Wyon, M. Parker and J. E. Gordon
1965 "A case study of birth interval dynamics", Population Studies, 19, 1 (July): 81-96.

Retel-Laurentin, A.
1973 "Fécondité et syphilis dans la région de la Volta Noire", Population 28, 4-5 (juillet-octobre): 793-815.

Sheps, M. C., J. A. Menken, J. C. Ridley and J. W. Lingner
1967 "Birth intervals: artifact and reality", Contributed papers, Sydney conference, IUSSP: 857-868.
1970 "Truncation effect in closed and open birth interval data", Journal of the American Statistical Association, 65: 678-693.

Sheps, M. C. and J. A. Menken
1972 "Distribution of birth intervals according to the sampling frame", Theoretical Population Biology, 3 (march): 1-26.
1973 Mathematical Models of Conception and Birth, University of Chicago Press, Chicago.

Tietze, C.
1960 "Probability of pregnancy resulting from a single unprotected coitus", Fertility and Sterility, 11 (September-October): 485-488.
1963 "The effect of breastfeeding on the rate of conception", International Population Conference, New York, 1961, Vol. II: 129-135, London.

Vincent, P.
1950 "La stérilité physiologique des populations", Population, 5, 1: 45-64.
1961 "Recherches sur la fécondité biologique", I.N.E.D.-P.U.F., Paris.

Warner, S. L.
1965 "Randomized response: A survey technique for eliminating evasive answer bias", Journal of the American Statistical Association 60: 63-69.

Wolfers, D.
1968 "The determinants of birth-intervals and their means", Population Studies 22, 2 (July): 253-262.

World Health Organization
1970 "Spontaneous and induced abortion", Technical Report Series, No. 461.

Wyon, J. B. and J. E. Gordon
1971 The Khanna Study. Population Problems in the Rural Punjab, Harvard University Press, Cambridge.

SOURCES FOR TABLES

(1) From J. W. Eaton and A. J. Mayer (1953), "The social biology of very high fertility among the Hutterites", Human Biology, 25: 206 *and* L. Henry (1961), "Some data on natural fertility", Eugenics Quarterly, 8: 81.

(2) Henripin, J. (1954), "La population canadienne au début du XVIII[e] siècle", P.U.F., Paris.

(3) Statistique générale du mouvement de la population 1749-1905 (1907), Statistiques générales de la France, Paris.

(4) Leridon, H., Zucker and M. Cazenave (1970), "Fécondité et famille en Martinique", I.N.E.D.-P.U.F., Paris *and* Leridon, H. (1971), "Les facteurs de la fécondité en Martinique", Population, 1971, No. 2.

(5) Potter, R. G., M. L. New, J. B. Wyon and J. E. Gordon (1965), "A fertility differential in eleven Punjab villages", Milbank Memorial Fund Quarterly 18; 185-201.

(6) Lorimer, F. *et al.* (1954), Culture and Human Fertility, UNESCO, Paris: 26.

(7) Cross, H. E. and V. A. McKusick (1970), "Amish Demography", Social Biology 17: 83-101.

(8) Cantrelle, P. (1969), "Etude démographique dans la région du Siné-Saloum (Sénégal)", O.R.S.T.O.M., Paris.

(9) Adapted from Hyrenius, H. (1958), "Fertility and reproduction in a Swedish population group without family limitation", Population Studies 12: 121-130.

(10) Gautier, E. and L. Henry (1958), "La population de Crulai, paroisse normande, I.N.E.D.-P.U.F., Paris: 97.

(11) Charbonneau, H. (1970), Tourouvre-au-Perche aux XVII^{e} et XVIII^{e} siècles, I.N.E.D.-P.U.F., Paris: 97.

(12) Fresel-Lezey, M. (1969), Histoire démographique d'un village en Béarn: Bilhères-d'Ossau, Pau: 103.

(13) Ganiage, J. (1963), Trois Villages de l'Ile de France", I.N.E.D.-P.U.F., Paris: 85.

(14) Valmary, P. (1965), "Familles paysannes au XVIII^{e} siècle en Bas-Quercy: Etude démographique", I.N.E.D.-P.U.F., Paris: 120.

(15) French, F. E. and J. E. Bierman (1962), "Probabilities of fetal mortality", Public Health Reports 77, No. 10.

(16) Taylor, W. F. (1970), "The probability of fetal death", in Congenital Malformations, Fraser and McKusick (eds.), Excerpta Medica, New York.

(17) Petterson, F. (1968), "Epidemiology of early pregnancy wastage: biological and social correlates of abortion", Svenska Bokforlaget, Norsteds, Stockholm.

(18) Shapiro, S., E. Jones and P. Densen (1962), "A life table of pregnancy terminations and correlates of fetal loss", Milbank Memorial Fund Quarterly XL (January).

(19) Erhardt, C. L. (1963), "Pregnancy losses in New York City", American Journal of Public Health, 53 (September).

(20) Cantrelle, P. and H. Leridon (1971), "Breast-feeding, mortality in childhood and fertility in a rural zone of Senegal", Population Studies 25, 3 (November): 505-533.

(21) Knodel, J. (1968), "Infant mortality and fertility in three Bavarian villages: an analysis of family histories from the 19th century", Population Studies 23, 3 (November): 297-318.

(22) Jain, A. K. (1969), "Pregnancy outcome and the time required for next conception", Population Studies 23, 3 (November): 421-433.

(23) Smith, T. E. (1960), "The Cocos-Keeling Islands. A demographic laboratory", Population Studies 14, 2 (November): 94-130.

Chapter 5
USE-EFFECTIVENESS OF CONTRACEPTION

by
Robert G. Potter and Roger C. Avery

I. INTRODUCTION

The previous chapters dealt with the acceptance of contraception. Later chapters take up births averted by contraception. The present chapter focuses on the problem of measuring use-effectiveness of contraception.

Broadly defined, use-effectiveness represents "the effect that might result from the use of one or more conception-regulating methods" (United Nations, 1969a). A much narrower view of the concept is adopted here in order to limit the discussion of measurement problems to a chapter of reasonable length.

Use-effectiveness of contraception has two basic dimensions: continuation and effectiveness. Effectiveness is the average extent to which natural fecundability, the woman's monthly chance of conception during a fecundable month, is reduced by use of the contraceptive. Continuation relates to the number of months she practices it. A perfectly effective contraceptive can offer but limited protection if practiced only for a few months. Given the high effectiveness of modern contraceptives, the continuation aspect becomes crucial.

Use-effectiveness is conventionally assessed by means of a multiple decrement life table methodology (United Nations, 1969a, b), an approach which has been shaped by four factors:

(1) To assess continuation, one needs to know the proportions of acceptors that are still successfully using the contraceptive at specified intervals after adoption.
(2) One wants to measure the relative intensities of the several causes for interrupting practice of the method, of which accidental pregnancy is only one.

(3) In the usual follow-up study there is a fixed cut-off date but variable dates of acceptance, resulting in records of unequal length and many respondents classified as continuing users at the end of the study. Sometimes contraceptive histories complete with length of use and cause of termination are available for only a minority of acceptors. To avoid bias and waste of information, one must incorporate the incomplete as well as the complete records into the analysis.
(4) As time elapses from acceptance, the monthly rates of termination, as exhibited by a cohort of acceptors, typically decrease at a decelerating rate. That the probability of continuing use for another month is not constant considerably complicates any attempt to extrapolate continuation beyond the period of observation. Yet such extrapolations are of prime importance when estimating births averted by contraception.

This chapter has three objectives:

(1) The first is to review the life-table methodology, paying attention to the manner in which it meets the four problems outlined above. For purposes of illustration, and more specifically to investigate the decline in the monthly risk of termination as a function of duration from acceptance, the method is applied to the fourth round of data from the Taichung Medical IUD Follow-up Study, distinctive for having records stretching back as long as nine years.
(2) Secondly, three techniques for extrapolating continuation beyond the period of observation are tested by means of the same Taichung data.
(3) A third purpose of the chapter is to address certain extensions of, and alternatives to, the life-table approach, which are appropriate when interests widen to include more than one segment of use of a single contraceptive or the interacting of several contraceptives.

II. LIFE-TABLE APPROACH

We consider the basic rationale of the life-table approach and how it meets the conditions defined earlier. Several more comprehensive expositions of the technique have been made (Freedman and Takeshita, 1969, Appendix 2; Potter, 1966 and 1967; Ross *et al.*, 1969; Tietze, 1967; Tietze and Lewit, 1968). To be concrete, let us assume that we are dealing with a set of first segments of IUD. Each segment is coded as to length and terminal status. Length is defined in terms

of ordinal months. Possible terminal statuses are pregnancy with device *in situ*, expulsion, removal, loss to follow-up or continuing use at cut-off date.

The following notation will be found convenient.

N_x	number of women retaining the device at start of the monthly interval $(x, x+1)$—i.e., the $(x+1)$th ordinal month
P_x	number of pregnancies during month $(x, x+1)$
E_x	corresponding number of expulsions
R_x	corresponding number of removals
$T_x = P_x + E_x + R_x$	total observed terminations during month $(x, x+1)$
F_x	number of women lost to follow-up during month $(x, x+1)$
C_x	number of continuing users last observed during month $(x, x+1)$
$W_x = C_x + F_x$	total women withdrawing from observation during month $(x, x+1)$

Two important simplifying assumptions are:

(1) that terminating events are uniformly distributed within an ordinal month and hence preceded on average by a duration of half a month.

(2) that women lost to follow-up or passing from observation because the cut-off date has occurred during ordinal month $(x, x+1)$ are no different in their risks of termination from those effectively observed over the entire month.

Let ordinal month $(m-1, m)$ represent the longest duration from insertion yielding enough cases N_{m-1} to support reasonably stable monthly rates. First computed are monthly rates of losing the device to a specific cause (pregnancy, expulsion, or removal) during the ordinal month $(x, x+1)$, for $x = 0, 1, \ldots, m-1$, given that the device was *in situ* at the start of the month.

$$Q_{xp} = P_x / N_x^* \tag{1}$$

$$Q_{xe} = E_x / N_x^* \tag{2}$$

$$Q_{xr} = R_x / N_x^* \tag{3}$$

where

$$N_x^* = N_x - W_x/2$$

The rationale underlying N_x^* is that had W_x device wearers not been lost to observation during month $(x, x+1)$, there would have been

$W_x/2$ additional months of exposure. Thus, for example, in computing the conditional monthly risk of pregnancy, one wants Q_{xp} to be the solution of

$$Q_{xp} = \frac{P_x + (W_x/2)\, Q_{xp}}{N_x} \tag{4}$$

which is satisfied by

$$Q_{xp} = P_x/N_x^*$$

That the three specific monthly risks of termination add up to the conditional monthly rate of losing the device for the three reasons combined is seen from

$$Q_{xp} + Q_{xe} + Q_{xr} = (P_x + E_x + R_x)/N_x^* \tag{5}$$

$$= T_x/N_x^* \tag{6}$$

$$= q_x \tag{7}$$

Conversely,

$$Q_{xp} = (P_x/T_x)\, q_x \tag{8}$$

and analogously for Q_{xe} and Q_{xr}. Finally, the conditional likelihood of surviving month $(x, x+1)$ as a continuing user is

$$p_x = 1 - q_x$$

All the above monthly rates are duration-specific and conditional on retention up to the start of the particular ordinal month.

The next computational step is to chain together appropriately these conditional monthly rates so as to obtain cumulative or monthly rates that relate to the entire cohort of acceptors instead of just the subsample retaining devices up to the start of a designated ordinal month. The proportion of the cohort expected to retain the device until the end of month $(x, x+1)$ is given by

$$P_{0\,(x+1)} = \prod_{i=0}^{x} p_i \tag{9}$$

and the cumulative proportion expected to terminate it sometime before end of month $(x, x+1)$ by

$$Q_{0\,(x+1)} = 1 - P_{0\,(x+1)} \tag{10}$$

The proportion of acceptors expected to terminate their device during the particular month $(x, x+1)$ is

$$d_x = P_{0x} - P_{0\,(x+1)} \tag{11}$$

$$= P_{0x}\, q_x \tag{12}$$

The corresponding proportions terminating the device during month $(x, x+1)$ for specific reasons are

$$d_{xp} = P_{0x} Q_{xp} \tag{13}$$

$$d_{xe} = P_{0x} Q_{xe} \tag{14}$$

$$d_{xr} = P_{0x} Q_{xr} \tag{15}$$

Summing these proportions for a particular risk yields the cumulative rates of device loss for that risk. For example, the cumulative rate of device loss because of pregnancy by the end of month $(x, x+1)$ is the sum

$$Q_{0(x+1)p} = \sum_{i=0}^{x} d_{ip} \tag{16}$$

With respect to our four requirements, the cumulative proportions P_{0x}, $x = 0, 1, \ldots, m$, are meeting our need for detailed information about continuation. Moreover, while the length of the follow-up period essentially dictates m, the maximum duration for which useful rates can be computed, it does not directly affect the values of rates obtained within this duration. In this sense, study results are independent of the length of observation—a condition which an earlier methodology based on the Pearl pregnancy rate did not fulfil (Tietze, 1962). In addition, multiple causes of discontinuation are being recognized, of which pregnancy is only one. Each incomplete history contributes to the exposure base N_x for as many months as are covered by the history. Finally, the conditional monthly rates Q_{xp}, Q_{xe}, and Q_{xr} are free to vary from month to month so that no simplifying assumption about their being constant has to be made.

Three extensions of the life-table approach are relevant to the applications to follow. All the specific rates considered so far have been "net" rates which allow for the presence of competing risks. For example, the net cumulative rate of pregnancy $Q_{0(x+1)p}$ allows for, and is reduced by, some women expelling the device or having it removed before they have had a chance to conceive accidentally. For certain purposes, such as investigating the declines in intensity of particular termination risks with duration from acceptance, it is desirable to measure the force of a particular termination risk independently of the competition of other types of termination. Women who discontinue during a month for reason of expulsion or removal are not exposed to the risk of pregnancy for a full month. From this perspective the net conditional monthly rate Q_{xp} underestimates the "gross" risk of pregnancy standing by itself. An estimate of the latter is

$$q_{xp} = P_x / N^{**}_{xp} \tag{17}$$

where

$$N^{**}_{xp} = N_x - (E_x + R_x + W_x)/2 \tag{18}$$

Note that the gross monthly rate of pregnancy during month $(x, x+1)$ satisfies the equation

$$q_{xp} = \frac{P_x + 0.5\,(E_x + R_x + W_x)\,q_{xp}}{N_x} \tag{19}$$

Let $p_{xp} = 1 - q_{xp}$. The monthly probabilities of escaping the gross risk of pregnancy may be chained to yield the likelihoods of avoiding, or succumbing to, the gross risk of pregnancy by the end of month $(x, x+1)$:

$$p_{0\,(x+1)\,p} = \prod_{i=0}^{x} p_{ip}, \tag{20}$$

$$q_{0\,(x+1)\,p} = 1 - p_{0\,(x+1)\,p} \tag{21}$$

Gross rates are not additive; the sum of monthly rates $q_{xp} + q_{xe} + q_{xr}$ exceeds q_x, the monthly rate of termination for all three reasons combined. Gross rates are also more hypothetical than net rates inasmuch as they involve an additional simplifying assumption, namely, that during any ordinal month $(x, x+1)$ the women observed to terminate from the several competing causes are unselected with respect to the single risk under review.

If the follow-up period is too long to make it desirable to present results by single months, one may reduce the amount of detail and increase the stability of reported rates by considering the conditional probabilities of continuing use or termination within the next k months if already having retained 0, k, 2 k, ... months. The formulae for k-month termination or retention given j months of previous retention are

$$Q_{0k/j} = 1 - \prod_{x=j}^{j+k-1} p_x \tag{22}$$

$$P_{0k/j} = \prod_{x=j}^{j+k-1} p_x \tag{23}$$

In the next section, recourse is made to conditional rates of 24-month retention and also to conditional gross rates of 6-month termination. For example, the gross risk of pregnancy in the next 6 months, given j previous months of retention, is

$$q_{06p/j} = 1 - \prod_{x=j}^{j+5} p_{xp} \tag{24}$$

Let R_{0x} stand for expected months of retention per acceptor by the end of month $(x-1, x)$. Then,

$$R_{0x} = \sum_{i=1}^{x} P_{0i} + 0.5\, Q_{0x} \qquad (25)$$

R_{0x} is unambiguously defined for $x = 0, 1, \ldots, m$. However, for estimating births averted by contraception, one may need an estimate of total retention. This means extrapolating beyond the observation period. Such would be easy if it could be assumed that after some duration from acceptance, $h < m$, the conditional monthly continuation rate p_x becomes constant, say equal to p. Then, if $q = 1 - p$,

$$R_{0\infty} = \sum_{i=1}^{h} (P_{0i} + 0.5\, p_{i-1}) + P_{0h} \sum_{i=1}^{\infty} (p^i + 0.5\, p^{i-1} q) \qquad (26)$$

$$= \sum_{i=1}^{h} P_{0i} + P_{0h} \left(\frac{p}{q}\right) + 0.5 \qquad (27)$$

If we cannot assume the constancy of p_x starting with some duration less than the effective observation period m, it is necessary to have a curve expressing the change in p_x for x-values beyond the observation period.

AN APPLICATION

The original sample of the Taichung Medical IUD Study consisted of women who had accepted an IUD in the Taichung intensive family planning program between July 1962 and May 1965 and who were residents of the city of Taichung at that time (Freedman and Takeshita, 1969). Whereas earlier analyses of these data have used cut-off dates of December 1964, December 1966, and July 1968 (B. Moots, 1972), the present analysis uses a cut-off date of December 1971.

Loss to follow-up had been relatively light, only 5.4 per cent of the original sample having been lost prior to termination of the first device. Of these, three quarters had moved out of the Taichung area and were not followed (cf. R. Avery, 1973 for more detail).

Annual contact had been maintained with most of the women. If an IUD was believed to be in place, the interview and an examination were made in the clinic; if there was no IUD, then the interview was conducted at home. Respondents who failed to keep clinic appointments were visited by nurses who attempted to get them to come to the clinic and failing that, interviewed them at home. The regimen of yearly contacts was suspended when it was ascertained that a woman had moved out of Taichung, had been sterilized, had reached menopause, or had become divorced, widowed, or separated from her husband.

Table 5.1. *Six-month termination probabilities, Taichung IUD-IV first segment gross rates*

Adjusted no. of women exposed	*Months since insertion*	*Proportion retaining to end of period*	*Six-month conditional gross rates of termination*					
			All causes	*Expulsion*	*Pregnancy*	*Medical removal*	*Removal to plan a pregnancy*	*Removal for other reasons*
5785	0	100.0	23.1	8.7	4.4	10.5	0.3	1.2
4553	6	76.9	16.8	4.6	4.4	6.5	1.4	0.9
3754	12	64.0	14.1	2.4	4.3	5.6	1.6	0.9
3182	18	55.0	13.4	1.4	4.4	5.3	2.0	0.9
2742	24	47.6	11.1	1.3	2.6	5.1	1.3	1.3
2381	30	42.3	10.8	1.0	2.0	5.6	1.1	1.6
2118	36	37.8	8.9	0.6	2.0	4.4	0.9	1.0
1899	42	34.4	6.4	0.5	1.2	3.6	0.7	0.5
1759	48	32.2	6.7	0.5	1.3	3.8	0.4	0.8
1623	54	30.0	7.3	0.5	1.1	4.2	0.5	1.2
1517	60	27.8	5.8	0.8	0.9	3.4	0.1	0.6
1398	66	26.2	5.6	0.4	0.8	3.1	0.1	1.1
1311 (1315)	72	24.8	4.6 (4.5)	0.1	0.1	3.0	0.0	1.3
1160 (1232)	78	23.6 (23.7)	4.1 (3.7)	0.1	0.0	2.9 (2.6)	0.1	1.0 (0.9)
946 (1021)	84	22.7 (22.8)	4.9 (4.3)	0.3 (0.2)	0.4	3.2 (2.8)	0.1	1.0 (0.9)
465 (541)	90	21.6 (21.8)	6.2 (5.3)	0.4 (0.3)	0.4	4.7 (3.9)	0.0	0.7 (0.5)
465 (541)	96	20.2 (20.6)	7.3 (4.9)	0.3 (0.2)	0.0	4.8 (3.1)	0.0	2.0 (1.4)

Numbers in parentheses are calculated assuming women still using at last contact in 1971 continued to the end of 1971.

With one exception, discussed below, lengths of first segments were defined conservatively, continuation being measured to date of last interview or to time of termination when known. As time since insertion increases, the number of women still exposed to the risks of termination decreases. Nevertheless, sample size remains large enough to yield adequately stable rates for the first eight years.

Table 5.1 provides cumulative rates of continuation and discontinuation for all cases, by six-month intervals, as well as gross six-month rates of termination for specific causes. The numbers in parentheses refer to a life table in which women still using a device at the last interview in 1971 were assumed to have continued to use it through the end of 1971. The other numbers assumed that women had continued to use the device only to the date of last interview. The former procedure has a slight downward bias on termination rates; some women will have terminated the device during 1971 but this fact will not have been reported. The latter procedure has an upward bias on termination rates; most terminations in the period of reference, being removals, will have occurred at the clinic and will have been recorded immediately. Those who have had an interview are more likely to have terminated than others, and since only women with an interview were included, the termination rates are biased upward. This problem only affects the termination rates at duration periods within a year of the cut-off date.

It can be seen that the conditional six-month probabilities of termination for all causes decline, at first precipitously and then more gradually, but do not approach anything like a plateau until after five years. Two explanations are available for this decelerating decline. First, women may share identical conditional monthly risks of termination with this shared risk declining as duration from insertion increases. Alternatively, each woman may have a characteristic conditional monthly risk of terminating which remains constant but varies among women. As time elapses from insertion, the more risk-prone tend to eliminate themselves by termination, leaving behind a subsample progressively selected for lower than average risk and reduced heterogeneity. One anticipates that both principles—that of duration-dependent change in individual risks as well as heterogeneity and selective elimination among women—are operative.

To investigate the matter more closely, one may examine the behavior of each source of termination. Use of gross rather than net rates helps to highlight contrasts. The behaviors of the expulsion and medical removal rates fit well the hypothesis of heterogeneity. The initial plateau of two years displayed by the pregnancy rate suggests a temporary opposition of forces, with the passage of women

from postpartum anovulation to fecundability tending to raise the pregnancy rate while the selective removal of pregnancy-prone women among the already fecundable women, lowers it. After two years the latter mechanism becomes dominant. "Removals to plan a pregnancy" constitute a minor attrition presumably because most of the Taichung women adopted an IUD in order to limit rather than to space births. If among the minority of spacers, preferred postponements of next pregnancy range mostly from 6 to 30 months, this would explain the rise in six-month rates to a peak during months 18-24 and then a subsequent decline essentially to zero. The conditional rates under the category "removal for other reasons" manifest no obvious trend.

The two specific risks subject to an important degree of duration dependency—namely pregnancy and removals to plan a pregnancy—have become minor after five years. Medical removals which, as remarked earlier, behave in a way consistent with the heterogeneity hypothesis, then become dominant.

Another means of investigating the decelerating decline of the

Table 5.2. *Probability of continuing IUD use during two-year intervals by selected characteristics, Taichung IUD-IV, first segment*

Characteristic	Number of women at duration 0	Interval			
		0-24 mo.	24-48 mo.	48-72 mo.	72-96 mo.
Age and reason for accepting					
Limiters <24	403	0.328	0.532	0.660	0.861
25-29	1,446	0.411	0.616	0.750	0.849
30-34	1,717	0.533	0.719	0.779	0.865
35-47	1,332	0.652	0.750	0.795	0.805
Spacers <24	423	0.176	0.319	0.500	0.625
25-29	354	0.234	0.351	0.703	0.652
30-34	58	0.404	0.571	0.656	0.686
Parity (live births and still births)					
0-1	193	0.132	0.277	0.401	—
2	690	0.291	0.483	0.702	0.766
3	1,126	0.379	0.604	0.736	0.791
4	1,214	0.476	0.663	0.756	0.869
5	956	0.532	0.702	0.778	0.830
6-17	1,606	0.627	0.745	0.793	0.838

six-month termination rate for all causes is by the study of differentials. Unless the sample is heterogeneous, one cannot expect that when subclassified by a variable such as age or parity it will exhibit pronounced differentials. Moreover, if through the mechanism of selective elimination of high-risk members, the sample is being made more and more homogeneous, then the magnitude of any set of differentials should shrink as duration increases.

Two dramatic confirmations of this expectation are found in Table 5.2 based on a classification of age of wife at time of insertion combined with reason for acceptance and parity. The rates are 24-month probabilities of retention. Both classifications generate very strong differentials during the initial 24-month period following insertion. However, by durations 48-72 months or 72-96 months, these differentials are much attenuated.

III. METHODS OF EXTRAPOLATION

To extrapolate P_{ox}, the proportion retaining, and R_{ox}, the expected months of retention per acceptor, beyond the observation period, in the face of sample heterogeneity, one must have either a model defining the conditional risks of termination as a function of duration since acceptance, or a mathematical function giving the direct relationship between proportions retaining and time elapsed.

Best known and most widely used among the latter is a continuous function based on the negative exponential proposed by Mauldin, Nortman, and Stephan (1967), namely.

$$P_{0x}^{M} = ae^{-rx} \tag{28}$$

where

a is a constant denoting the proportion who avoid very early termination,

r is a constant reflecting the annual rate of termination, and

e is the natural log base 2.718 .... Integrating this decay function over duration, one has

$$R_{0x}^{M} = \int_{0}^{x} ae^{-rt}dt \tag{29}$$

$$= \frac{a}{r}(1 - e^{-rx}) \tag{30}$$

Note that as $x \rightarrow \infty$,

$$R_{0x}^{M} \longrightarrow a/r$$

This model has proved to be quite a good fit for IUD experiences

(including the first two rounds of the Taichung IUD Follow-up Study) which exhibit an initial rapid decline of conditional monthly termination rates followed by a plateauing of risk. The level of the plateau is measured by the parameter r while the parameter a allows for the relatively higher loss rates of the first few months. However, in the case of the Taichung Follow-up data, the upward bias in termination rates for the last year of observation, mentioned in connection with Table 3.1, may have contributed to the appearance of a levelling off when the maximum duration m for computation was still as short as 24 or 36 months.

Be that as it may, a temporary decline of conditional risks followed by a levelling off within a year could be expected from a sample of acceptors composed of two types: a minority of high-risk couples most of whom are eliminated by termination within a year and a majority who because of much lower risks become dominant after a year so that thenceforth the sample exhibits an essentially constant conditional rate of device loss.

Recently Liu *et al.* (1972) have presented a model which formalizes the situation just described. In their three-parameter model, the high-risk group has a constant conditional risk of termination r_1 (per year); the low-risk group has a constant conditional risk of r_2; while a proportion p of the initial cohort belong to the high-risk group and $1-p$ belong to the low risk group. With the process thus represented as a mixture of two negative exponentials, operating in continuous time,

$$P^{L}_{0x} = pe^{-r_1 x} + (1-p)\,e^{-r_2 x} \tag{31}$$

$$R^{L}_{0x} = \frac{p}{r_1}(1-e^{-r_1 x}) + \frac{(1-p)}{r_2}(1-e^{-r_2 x}) \tag{32}$$

Both the Mauldin *et al.* and Liu *et al.* functions are conventionally fitted by a least squares procedure (after first passing to logarithms) applied to the monthly proportions retaining, $P_{01}, P_{02}, \ldots, P_{0m}$, derived from a life table analysis. This least squares procedure is well described in the article by Liu *et al.* (1972).

The model of Liu *et al.* is too new to have been widely applied. Because it possesses three parameters instead of two, it will presumably give slightly better fits than those of the Mauldin *et al.* model in the case of the contraceptive experiences for which both are designed, namely, those exhibiting an initial decline of conditional loss rate followed by a leveling-off, all within the observation period. Within this class of experiences, the slower the decline, the greater the improvement that the Liu *et al.* fitting can be expected to produce over the Mauldin *et al.* fitting. However, when the observation period

is too short to encompass the plateauing of conditional risk of device loss and when the decline continues past the end of the observation period, then both procedures—the mixture of two negative exponential as well as the single negative exponential—can be counted upon to produce conservative results.

This conservatism is demonstrable with Taichung IUD-IV data. One can pretend that one has available only the first 12, 24, 36, . . ., up to 96 months of experience and in each instance fit the two functions in order to see how sensitive their results are to length of observation.

From Table 5.3, it is evident that the parameters a and r as estimated by the Mauldin *et al.* procedure are highly sensitive to length of observation. The same applies to parameters p, r_1, and r_2 obtained by the Liu *et al.* technique.

able 5.3. *Results of fitting the methods of Mauldin* et al. *and Liu* et al. *to differing lengths of IUD-IV experience*

	Method of Mauldin et al.			*Method of Liu* et al.			
					Monthly discontinuation risk		
Length of observation (months)	*Proportion not immediately discontinuing* a	*Monthly risk of discontinuing* r	*Retention during (0,96)* $R_0^{M}{}_{(96)}$	*Proportion in high risk group* p	*high* r_1	*low* r_2	*Retention during (0,96)* $R_0^{L}{}_{(96)}$
12	0.95	0.034	27.5	0.08	1.060	0.031	28.7
24	0.92	0.028	30.3	0.12	0.323	0.026	31.5
36	0.89	0.025	32.2	0.15	0.237	0.023	33.2
48	0.85	0.022	33.8	0.20	0.186	0.020	34.9
60	0.81	0.019	35.2	0.24	0.155	0.018	36.3
72	0.78	0.018	36.0	0.27	0.137	0.016	37.3
84	0.74	0.016	36.6	0.31	0.124	0.015	38.1
96	0.71	0.014	36.1	0.34	0.115	0.013	38.8

From the life table analysis, the estimate of mean retention per woman during interval (0, 96), based on

$$R_{0\,(96)} = \sum_{x=1}^{96} P_{0x} + 0.5\, Q_{0\,(96)} \qquad (33)$$

is 37.5. As hypothesized, both techniques tend to underestimate mean retention time, the Mauldin *et al.* method slightly more than the Liu *et al.* method, and both more markedly when the observation period is short. The underestimate falls below 10 per cent at an observation length between 36 and 48 months for the Liu *et al.*

technique and between 48 and 60 months for the Mauldin *et al.* procedure (cf. Table 5.4). At longer durations (84 and 96 months) the Liu *et al.* technique overestimates retention.

Table 5.4. *Results of fitting the methods of Mauldin* et al. *and Liu* et al. *to differing lengths of IUD experience*

Length of observation (months)	*Method of Mauldin* et al.			*Method of Liu* et al.		
	Percentage underestimate of retention during (0,96) using:		*Theoretical vs. "observed"*	*Percentage underestimate of retention during (0,96) using:*		*Theoretical vs. "observed"*
	$R^{M}_{0(96)}$	$R^{*M}_{0(96)}$	Δ_M	$R^{L}_{0(96)}$	$R^{*L}_{0(96)}$	Δ_L
12	27	27	10.42	24	23	8.91
24	19	18	7.48	16	16	6.23
36	14	12	5.76	11	11	4.79
48	10	7	4.51	7	6	3.57
60	6	4	3.70	3	3	2.66
72	4	2	3.29	0.4	1.4	2.14
84	2	0.3	2.21	−2	0.3	1.88
96	2	—	3.39	−4	—	1.83

The prediction of $R_{0(96)}$ can be improved by using "observed" results (as taken from the life table analysis) for the length of the observation period m and extrapolating only over the remainder of the interval (m, 96). These hybrid formulae are:

$$R^{*M}_{0(96)} = \sum_{i=1}^{m} P_{0i} + 0.5\,Q_{0m} + P_{0m}\left(\frac{1}{r}\right)(1.0 - e^{-(96-m)\,r}) \qquad (34)$$

$$R^{*L}_{0(96)} = \sum_{i=1}^{m} P_{0i} + 0.5\,Q_{0m} + L\left(\frac{1}{r_2}\right)(1.0 - e^{-(96-m)\,r_2}) + H\left(\frac{1}{r_1}\right)(1.0 - e^{-(96-m)r_1}) \qquad (35)$$

where m is the length of the observation period and the proportions of the original cohort expected to be low- or high-risk and still retaining at duration m are given by

$$L = P_{0m}\left(\frac{(1-p)e^{-mr_2}}{(1-p)e^{-mr_2}+pe^{-mr_1}}\right) \tag{36}$$

$$H = P_{0m}\left(\frac{pe^{-mr_1}}{(1-p)e^{-mr_2}+pe^{-mr_1}}\right) \tag{37}$$

It may be recalled that p is the proportion of the original cohort who belong to the high-risk group.

The improvements of

$$R^{*M}_{0(96)} \text{ over } R^{M}_{0(96)} \text{ or } R^{*L}_{0(96)} \text{ over } R^{L}_{0(96)}$$

prove negligible (Table 5.4). Perhaps they are so small when observation is short, 12 or 24 months, because one is still having to predict over most of the interval (0, 96). On the other hand, when the observation period is 60 months or more, prediction by

$$P^{M}_{0m} \text{ or } P^{L}_{0m}$$

has become so efficient as to leave little room for improvement.

A more searching test of the relative goodness of fit of the two models is gained by summing the absolute differences of observed and predicted monthly proportions retaining, i.e.,

$$\Delta_M = \sum_{x=1}^{96} |P_{0x} - P^{M}_{0x}| \tag{38}$$

$$\Delta_L = \sum_{x=1}^{96} |P_{0x} - P^{L}_{0x}| \tag{39}$$

As Table 5.4 shows, in application to the Taichung IUD-IV data at least, the advantage consistently lies with the Liu *et al.* procedure.

Plainly the performance of the Liu *et al.* and Mauldin *et al.* functions is not so perfect as to remove all incentive for exploring a third approach. Instead of a mixture of two homogeneous populations as in the Liu *et al.* model, a continuous probability distribution of termination risks may be ascribed to the cohort of acceptors.

Our simplifying assumptions are:

(1) each couple is subject to a continuous *fixed* risk of discontinuation per unit of time,
(2) those fixed risks vary among acceptors as a Pearson Type III curve,
(3) discontinuation is a random event,
(4) the sample of experience is large.

The two-parameter density of discontinuation risks is

$$f(E) = \frac{a^k}{\Gamma(k)} e^{-aE} E^{k-1} \qquad a,\ k \geqslant 0,\ E \geqslant 0 \tag{40}$$

The density becomes L-shaped when $k<1$. The mean of discontinuation risks is given by k/a, the variance by k/a^2. Taken over the density $f(E)$, the proportion Q^T_{0t} expected to discontinue by time t is

$$\int_0^\infty f(E)\,(1-e^{-Et})\,dE = 1-\left(\frac{a}{a+t}\right)^k \tag{41}$$

The proportion P^T_{0t} expected to continue to duration t is the complement, or

$$\left(\frac{a}{a+t}\right)^k$$

a simple two-parameter function of duration since acceptance. Thus, we are assured that the proportion retaining will decline at a decelerating rate per unit of duration. Expected retention per acceptor during interval $(0, t)$ is

$$R^T_{0t} = \int_0^t \left(\frac{a}{a+x}\right)^k dx \tag{42}$$

$$= \frac{a^k}{1-k}\left[(a+t)^{1-k}-(a)^{1-k}\right] \tag{43}$$

The latter expression increases beyond all bounds with t when $k<1$, though giving meaningful values for finite t. For the case $k>1$, we use

$$R^T_{0t} = \frac{a^k}{k-1}\left[\left(\frac{1}{a}\right)^{k-1} - \left(\frac{1}{a+t}\right)^{k-1}\right] \tag{44}$$

which plainly increases with t to $a/(k-1)$ in the limit.

Only the heterogeneity principle is operative in this third model with change over time being dependent on the selective removal of the more risk-prone acceptors. No provision is being made for duration-dependent changes in the termination risks of individual women.

Although wasteful of information, the following fitting procedure suggested by R. N. Singh (Potter *et al.*, 1962) is convenient. Define $Q^T_{t,b}$ to be the proportion of acceptors who are expected to discontinue between durations a and b. Let

$$C = Q^T_{0,m/2}$$

and

$$C' = Q^T_{m/2,m}$$

where

m is the length of observation period, or more exactly the maximum duration for which the life table analysis produces reasonably stable rates. Equating the proportions C and C′ with their theoretical counterparts from the model,

$$C = 1 - \left(\frac{a}{a+m/2}\right)^k \tag{45}$$

$$C + C' = 1 - \left(\frac{a}{a+m}\right)^k \tag{46}$$

Forming a ratio of these two equations, after first passing to logarithms, one has

$$\frac{\log\left(\frac{a}{a+m/2}\right)}{\log\left(\frac{a}{a+m}\right)} = \frac{\log(1-C)}{\log(1-C-C')} \tag{47}$$

an equation in the single variable a, with k eliminated. In this last equation, a is solved by trial-and-error and then k obtained by

$$k = \frac{\log(1-C-C')}{\log\left(\frac{a}{a+m}\right)} \tag{48}$$

A test of the fit of the model to Taichung IUD-IV data is afforded by

$$\Delta_T = \sum_{x=1}^{96} |P_{0x} - P_{0x}^T| \tag{49}$$

Table 5.5. *Results of fitting type III-exponential distribution to initial* m *months of IUD-IV experience,* m = 12, 24, ..., 96

Length of observation (months)	*Parameters of distribution*		*Retention during (0,96)*	*Mean risk* k/a	*Standard deviation of risks*	
	a	k	$R_{0(96)}^T$		k/a^2	Δ_T
12	9.1975	0.5342	41.7	0.0581	0.0794	4.29
24	15.2812	0.7862	37.8	0.0514	0.0583	0.52
36	20.5312	0.9615	36.9	0.0468	0.0479	0.91
48	19.8750	0.9223	37.5	0.0464	0.0479	0.61
60	18.9257	0.8951	37.6	0.0473	0.0500	0.49
72	17.6625	0.8587	37.6	0.0486	0.0520	0.37
84	16.6687	0.8225	37.9	0.0493	0.0548	0.47
96	16.3500	0.8186	37.7	0.0501	0.0557	0.39

Three points are made clear by comparing Table 5.5 with Tables 5.3 and 5.4. The estimated parameters a and k of the Type III-exponential function are far less sensitive to length of observation period than are the parameters derived from either the Mauldin *et al.* or the Liu *et al.* techniques. Secondly, even at short durations the predictions of $R_{0(96)}$ by $R^{T}_{0(96)}$ are fairly adequate, being 11 per cent high at 12 months, but deviating less than 5 per cent from 37.5 months for any observation period 24 months or longer. Third, values of Δ_T are considerably below their Δ_M or Δ_L counterparts, indicating a better overall fit.

One limitation of the Type III-exponential function should be recalled. When $k<1$, which is usually the case in the present application, R_{0t} is defined only for finite values of t. Unlike with the Mauldin *et al.* and Liu *et al.* techniques, there is no unique value to assign $R_{0\infty}$. Rather one must adopt some maximum duration—for instance, the duration that would bring ages to an average of 45 years—with this rather arbitrary choice having a large bearing on the figure calculated for total retention.

IV. EXTENSIONS

When interest extends beyond single segments of a contraceptive, extensions or alternatives to the life-table approach described above become necessary. To keep an already long chapter within bounds, only a few of these possible extensions can be reviewed, notably reinsertion rates, multiple segments, extended use-effectiveness, and for the treatment of several contraceptives over multiple pregnancy intervals, Markov chains. A basic consideration is the degree of independence existing between the index used and the length of the observation period. A failure to provide this independence may be taken as a serious flaw indeed.

The life-table approach described above is most easily and efficiently applied to first segments of IUD. In the case of oral contraception (cf. Balakrishnan *et al.*, 1970) and conventional contraceptives, the uncertainties of whether supplies are used and the possibility of irregular practice make the definition of segment length difficult. These problems are discussed by Chandrasekaran and Karkal (1972) in their application of the life-table approach to retrospective data pertaining to the diaphragm and jelly method.

A) *REINSERTION RATE*

Because so many women who terminate IUDs have another

device inserted, reinsertion rates are a significant item in the total picture of IUD continuation. However, the probability of one or more reinsertions is partly a function of duration. Suppose that in an analysis of follow-up, duration is being measured from first insertion. A woman who retains her first device for t months has no chance of reinserting until the end of that period and then, should she fail to accept another device at the time of having the first one removed, she has an increasing chance of later reinsertion the longer the observation period endures beyond t. Thus, the reinsertion rate of a cohort of IUD acceptors is correlated inversely with first segment continuation and directly with length of observation. To escape these dependencies and attendant disadvantages, one may appropriately measure duration from termination of first device and report the proportions having a first reinsertion within specified intervals after the first termination.

B) *MULTIPLE SEGMENT*

Faced with the irregular practice of oral or conventional methods of contraception, one may decide to ignore brief interruptions of usage provided that no pregnancy intervenes, and is thus led to "multiple segments" as a unit of analysis. Even in the case of the IUD, if a woman expels or has the device removed but then has a prompt reinsertion, it seems proper when measuring method performance to pool the two or more segments to form a multiple segment whose termination is that of the last segment. The multiple segments may be treated in a life table analysis in exactly the same manner as described above for first segments. Indeed, there exists an extensive literature of IUD continuance based on multiple segments.

Problems arise with the multiple segment concept when non-trivial periods of nonuse between individual segments are common. It is disputable whether or not to include these interregnums of nonuse as part of segment length. Moreover, regardless of whether they are or are not included, there remains an unwanted dependence upon length of follow-up period. In the case of IUDs, if one waits long enough, then a few more devices will be reinserted and if between prior and presently reinserted device, no pregnancy occurs, then the present IUD (and perhaps the interval of nonuse) are added to a previous segment length.

C) *EXTENDED USE-EFFECTIVENESS*

There may be not only irregular practice of the coitus-related methods, but frequent switching back and forth among several

methods. Rather than try to keep track of so many segments and the interspersed periods of nonuse, one may prefer to concentrate on the timing of the next conception. The question then becomes: among a group of women wishing to postpone pregnancy, how many are succeeding after a duration of 1, 2, . . . m months. One no longer keeps track of the length of contraceptive practice.

However, the problem arises of what to do about those who deliberately interrupt contraception in order to conceive or who no longer have need for contraception by virtue of a divorce, sterilization, etc. In the useful and subtle notion of extended use-effectiveness, developed by Tietze and Lewit (1968), the experience of such women is utilized as part of the denominator N_x up to the time of deliberately suspending efforts to postpone pregnancy or ceasing to need contraception. The life table analysis is organized to yield gross rates of pregnancy. Segment length is measured as the interval from a defined origin (previous pregnancy termination, clinic visit, acceptance of a designated method, etc.), to next conception, suspension of efforts to postpone pregnancy, cessation of need, cut-off date, or loss to follow-up, whichever occurs soonest. Appropriate to a gross risk, the denominator N_{xp}^{**} for computing a conditional monthly rate must register half a month for any women lost during ordinal month $(x, x+1)$ except by pregnancy.

The approach works better in a prospective study than a retrospective study, there being less opportunity to rationalize an unintended pregnancy as one happening immediately after deliberate interruption of contraception. It is well documented that spacers are subject to higher discontinuation rates (of any contraceptive) than are limiters. Hence, other factors being constant, the more the spacers and the fewer the limiters in a sample, the lower the extended use-effectiveness is likely to be.

The approach can be redefined to make the next live birth the terminating event if it is desired to encompass the impact of induced abortion as well as contraception. A worked example of this option is offered by Ross *et al.* (1969).

A strength of the extended use-effectiveness approach is that, as a variant of the life table analysis described earlier, it preserves an essential independence between study results and length of observation period m. Another strength of the extended use-effectiveness concept is that it transcends the complexity of method shifts and interruptions of contraceptive use, thereby retrieving simplicity and feasibility of analysis. This strength is also its weakness. By transcending complexity, it sacrifices detail and wastes information. One learns nothing directly about the acceptability and use-effectiveness

of the contraceptives that produce the derived distribution of intervals to next conception (or live birth).

V. MARKOV CHAINS

In the Taichung IUD Study, respondents have been followed even after they terminate a device and fail to have a prompt reinsertion. Information is available about the timing of method switches and lengths of use as well as about the timing, length, and outcome of pregnancies following the first IUD insertion. A continuing challenge has been how to analyze such rich data in a way that elucidates the interactions of several contraceptives with sterilization, induced abortion, and fertility, without being excessively wasteful of information or subverting independence between results and length of observation period.

A first step, taken by Freedman and his colleagues (Freedman and Takeshita, 1969; p. 304), was to classify respondents by a fixed set of states at specific durations from either initial insertion or termination of first device. These sets of states have indicated current status (using original IUD, subsequent IUD, other contraception, sterilized, not practicing any method, current pregnancy, etc.) and have sometimes been elaborated to give current status cross-classified by whether a pregnancy or birth has taken place since first insertion. The strengths and weaknesses of this approach are in some respects similar to those of the concept of extended use-effectiveness: the approach assures simplicity and feasibility of analysis but transcends, and therefore fails to illuminate, the relative acceptabilities, continuation, and effectiveness of the several interacting contraceptives that have helped to produce the distribution of respondents among states. The technique describes an end result in an orderly manner but does not explain how the result came about.

To meet the challenge, one of the present writers (R. A.) has been experimenting with a Markov chain approach. A time unit of v months is first selected—which might be six or perhaps better three months. Next a set of s mutually exclusive states is defined that are exhaustive in the sense that every respondent can be classified by one of them and no respondent fits two or more simultaneously. Each respondent is then classified as to state both at the start and finish of each v-month period for as many periods as her follow-up history covers. (If she is lost to observation during a period, that fragment of experience lasting $v-1$ months or less is not utilized in the analysis.) The pair combinations of states at start and finish of an interval define the logically possible transitions, s^2 in number.

Table 5.6. *Six-month probabilities of staying in same state* p_k (i, i), *by subsample, duration* k, *and state* i: *Taichung IUD-3 data*

Duration	IUD	Other contraceptive	No contraception	Sterilization
	Spacers			
(*a*)				
0-6	0.717	—	—	—
6-24	0.725	0.748	0.883	1.000
24-36	0.800	0.724	0.818	1.000
36-60	0.859	0.807	0.778	1.000
	Limiters ⩽ 29 years			
0-6	0.788	—	—	—
6-24	0.842	0.802	0.772	1.000
24-36	0.882	0.824	0.731	1.000
36-60	0.904	0.842	0.742	1.000
	Limiters 30-34 years			
0-6	0.860	—	—	—
6-24	0.894	0.822	0.759	1.000
24-36	0.919	0.831	0.780	1.000
36-60	0.943	0.870	0.789	1.000
	Limiters 35 and over			
0-6	0.871	—	—	—
6-24	0.924	0.867	0.803	1.000
24-36	0.928	0.874	0.851	1.000
36-60	0.939	0.847	0.884	1.000

(*a*) All intervals extend up to, but do not include, the end month shown.

To be clear about how transition probabilities are computed and subsequently chained, some further notation is useful. The kth interval spans duration $(k-1, k)$, duration being measured in units of v months from date of first insertion. Suppose that sample size and observation period are such as to permit the computation of reasonably stable transition probabilities over T intervals. Let the subscript k $(k = 0, 1, \ldots, T)$ signify duration in v-month units while subscripts i, j $(i, j = 1, 2, \ldots, s)$ identify states at start and finish of a v-month interval, respectively. Further, let N_{k-1} denote the number of women who are observed to terminate during the kth interval or to continue practice through the interval. Among these N_{k-1} who contribute experience relevant to the calculation of transition probabilities during the kth interval, $N_{k-1}(i)$ start the interval in state i while $N_{k-1}(i, j)$ both start in state i and finish the interval in state j. The transition probability $p_{k-1}(i, j)$ of shifting from state i to state j during the kth interval is thus defined by $N_{k-1}(i, j)/N_{k-1}(i)$. Properly,

$$\sum_{j=1}^{s} p_{k-1}(i, j) = 1$$

for all k. If for a particular state i, the following relations hold,

$$p_k(i, i) = 1.0$$

and

$$p_k(i, j) = 0, \quad i \neq j, \quad \text{for all } k,$$

then we may speak of state i as "absorbing", i.e., women can enter state i from another state but may not leave it for another. Other states are "transitional". Let M_{k-1} denote the $s \times s$ matrix of transition probabilities $p_{k-1}(i, j)$ pertaining to the kth time interval. Define $p_k(i)$ as the proportion of respondents who at duration k are in state i and P_k as the row vector whose elements are

$$p_k(i), \quad i = 1, \ldots, s.$$

Respecting the initial vector P_0, its elements are given by

$$p_0(i) = N_0(i)/N_0, \quad i = 1, \ldots, s.$$

That is, among the N_0 women observed throughout the initial interval, $p_0(i)$ is the proportion observed to be in state i at the start of the interval. Actually, since a cohort of IUD acceptors is being followed, the only nonempty state at duration 0 is that in which the original IUD device is retained.

For $k = 1, 2, \ldots, T$, we have

$$P_k = P_{k-1} M_{k-1}. \tag{50}$$

The expected distribution among states at duration T is

$$P_T = P_0 M_0 M_1 \ldots M_{T-1}. \tag{51}$$

If after some duration $k = h$, $h \leqslant T$, we may assume that the matrix of transition probabilities has become constant and equals M, say, then

$$P_T = P_0 \left(\prod_{k=0}^{h-1} M_k\right) M^{T-h} \tag{52}$$

$$= P_h M^{T-h} \tag{53}$$

During any interval, some respondents are lost to observation for such reasons as moves away, menopause, widowhood, etc. Hence, typically $N_k < N_{k-1}$.

The high discontinuation rates calculated for IUDs based on the first two rounds of the Taichung Follow-up Study, using multiple segments, would lead one to suspect very low proportions of the original acceptors still retaining a device after five years. Yet, the third and fourth rounds of Taichung data have been showing something quite different: the proportion of original acceptors currently wearing an IUD (including the Ota ring) has leveled off at around 50 per cent (Moots, 1972). One explanation is that many terminators after shifting to another method or after an appreciable period of non-practice or after one or more pregnancies return to the IUD. Direct evidence for this supposition is derived from a simple Markov chain.

The data for this chain are taken from the third round of the Taichung IUD Follow-up Study. Only four states are distinguished: IUD (including the Ota ring), other contraceptives, no contraceptive (i.e., non-practice of contraception, including being currently pregnant), and sterilization. Moves away, other loss to follow-up, menopause, death or widowing of respondent, divorce or separation, are all treated as withdrawals from observation (which decrease the size of N_{k-1} for the kth interval in which they occur). Calculations are extended over ten 6-month intervals. In our notation, $v = 6$, $s = 4$, $T = 10$, and M is 4×4.

For added interest, the sample is partitioned into four subsamples according to wife's age at time of first insertion and whether her goal at the time was spacing or limiting births. The four subsamples are: spacers of all ages, limiters of 29 years or less, limiters aged 30-34, and limiters of 35 and over. A Markov chain analysis is done upon each subsample. For the initial interval, 0-6 months, transition probabilities are calculated in a straightforward way. Over intervals 6-12, 12-18, and 18-24, transition probabilities are treated as constant and calculated by averaging the experience of the three intervals. In a similar manner, transition probabilities are viewed as constant

and averaged over the two intervals 24-30 and 30-36 and then again over the last four intervals 36-42, . . ., 54-60. In this fashion most of the duration-dependency is registered while the pooling of experience over two or more intervals increases the stability of estimated probabilities, especially for the longest durations.

In Table 5.6, six-month probabilities of staying in the same state are given. In all four subsamples these probabilities $p_k(1, 1)$ for IUD tend to increase with advancing duration k. Part of the explanation is the selective process toward progressively lower risk of termination among respondents continuing to wear the initial device. Then too, reacceptors may tend to be older and more motivated than first acceptors, thereby carrying lower risks of termination than when they were wearing their first device. This second mechanism would particularly apply to the spacer subsample, many of whom become limiters within 60 months of first insertion.

With some exceptions the probabilities of remaining in the "other contraceptive" category for six months also rise with increasing duration from first insertion. The trend in probabilities of remaining a nonuser for six months as related to duration varies depending on the subsample. Among women who have had a first insertion as

Table 5.7. *Six-month probabilities of shifting to, or away from, IUD, by method and subsample, during durations 6-24 and 36-60 months from first insertion: Taichung IUD-3 data*

Duration and subsample	Other contraceptive		No contraception		Sterilization	
	to IUD	from IUD	to IUD	from IUD	to IUD	from IUD
Duration 6-24 and:						
spacers	*0.007*	*0.031*	*0.068*	*0.243*	0.000	0.002
limiters:						
⩽29	0.038	0.033	*0.111*	*0.117*	0.000	0.008
30-34	0.025	0.023	0.111	0.077	0.000	0.006
35+over	0.031	0.013	0.101	0.060	0.000	0.002
Duration 36-60 and:						
spacers	0.025	0.021	0.133	0.117	0.000	0.003
limiters:						
⩽29	*0.026*	*0.033*	0.118	0.060	0.000	0.002
30-34	0.033	0.017	0.085	0.037	0.000	0.003
35+over	0.017	0.012	0.057	0.048	0.000	0.001

Underlined figures identify cases in which shifts away exceed shifts to IUD.

a spacer, the trend is downwards, suggesting a progressive shift from casualness about contraception as a spacer to more readiness to accept and continue contraception as a limiter. No clear trend is seen among limiters under 30 years and only a weak trend among limiters 30-34 years old when first accepting an IUD. Only amongst the oldest group of limiters does the probability of remaining a nonuser another six months definitely increase with lengthening duration. Perhaps with longer duration, a growing proportion are discovering a condition of subfecundity.

More directly relevant to the inquiry posed is Table 5.7. In each subsample, for two durations, rates of outflow from IUD are compared with rates of inflow. With respect to sterilization, an absorbing state, the shifts are all one way, away from the IUD. However, the absolute size of these rates of outflow (always less than .01) is small. With respect to "other contraceptives", or nonpractice of any method at all, the balance of shifts is favorable to IUD use in five out of eight cases for duration 6-24 months and in seven out of eight cases for duration 36-60 months. While outflow is substantial, inflow more than compensates. As was hypothesized, the long-term acceptability of the IUD (including the Ota ring) is high enough for it to recoup its losses to other methods or to nonuse of contraception by subsequent reacceptances.

The above limited application hardly does justification to the flexibility of the Markov chain approach. One has broad latitude in defining sets of states. For example, one might distinguish a larger number of contraceptives, also the state of current pregnancy. If postpartum anovulation is of concern, one can classify contraceptive status by interval since last live birth, e.g., 6 months or less, 7-12 months, 13 months or more. One is also free to choose which events—such as sterilization, menopause, divorce—are to be considered as withdrawals from observation (thereby tending to reduce N_k) and which as absorbing states (thereby adding to the number of state s).

A number of important parallels exist between the life table and Markov chain techniques. The existence of s states in the latter is analogous to the multiple causes of termination of the multiple decrement version of the former. Transition probabilities play a similar role to the conditional monthly probabilities in that they may be chained together, though the end result is proportionate distributions among states rather than cumulative rates. Like the conditional monthly probabilities, the transition probabilities typically change from one duration to another and therefore do not require an assumption of constancy. On the other hand, when an assumption of fixity becomes tenable, one gains a ready means of extrapolating beyond

the observation period. In both techniques, histories that are truncated by an early withdrawal from observation are utilized for as many time units as they qualify. Results from a Markov chain are duration-specific, thereby assuring the same essential independence between outcomes and length of observation as vouchsafed by a life table. (There is, of course, the qualification—to which the life table approach is subject, too—that if probabilities are changing over calendar time, then as a second order effect, the same length of observation period but starting at two different dates, or merely an extension of the first observation period, will lead to slightly different duration-specific results).

Finally, there exists a counterpart to the measurement of continuation in a life table. For any absorbing state i, one can calculate the proportion of the sample entering that state during any time interval $(k-1, k)$ merely by taking the first difference $p_k(i) - p_{k-1}(i)$. Moreover, if one is interested in something like the time to first pregnancy after the initial IUD insertion—a single decrement problem—one has only to redefine the transitional state "currently pregnant" (call it state j) into an absorbing state and examine the first differences $p_k(j) - p_{k-1}(j)$. To convert j into an absorbing state, set $p_k(j, j) = 1.0$ for all k and $p_k(j, i) = 0$ for all k and $i \neq j$ while leaving all other transition probabilities unchanged.

The same stratagem may be extended to the multiple decrement case. Suppose, for example, that in the application given earlier, one wished to trace the time of first switch to another method. One converts the transitional state "other contraceptive" into a second absorbing state, "sterilization" being already an absorbing state. Denote these two states as "2" and "4". The first differences $p_k(2) - p_{k-1}(2)$ and $p_k(4) - p_{k-1}(4)$ would provide the proportions of the sample switched from IUD (or possibly from nonuse of contraception) during the kth interval to the two competing methods.

Naturally there are limitations in applying the Markov chain approach to contraception. Few follow-up studies are able to provide the requisite data. Choice of time interval—for example, 3 months instead of 6 months—does affect results. With too long an interval, one misses events (e.g., pregnancies ending in abortion) and underestimates the rate of transitions. Too short an interval increases exponentially the work of coding and calculation without necessarily yielding a commensurate gain in information. Unless sample size is large, one is restricted as to the number of states that may be distinguished while still retaining usefully stable estimates of transition probabilities. Given the newness of the technique in a contraceptive setting, a versatile set of standard errors or tested rules of

thumb about optimal choice of time interval and advantageous classifications of states is still lacking. Obviously much work lies ahead before there can be a clear idea of the potential utility of this approach as an extension and generalization of the life table technique.

REFERENCES

Avery, R. C.

1973 "Patterns and implications of long-term IUD use in Taichung, Taiwan." Studies in Family Planning, 4 (June): 139-143.

Balakrishnan, T. R., *et al.*

1970 "Analysis of oral contraceptive use through multiple decrement life-table techniques." Paper submitted to 1970 meeting of the Population Association of America.

Chandrasekaran, C. and M. Karkal

1972 " 'Continuation rate,' 'use-effectiveness,' and their assessment for the diaphragm and jelly method," Population Studies 26 (November): 487-494.

Freedman, Ronald and John Y. Takeshita

1969 Family Planning in Taiwan. Princeton: Princeton University Press. Appendix X-2.

Liu, P. T., *et al.*

1972 "A Study of IUD Retention by Curve-fitting," Demography 9 (February): 1-11.

Mauldin, W. P., *et al.*

1967 "Retention of IUDs: an international comparison," Studies in Family Planning 18 (April): 1-12.

Moots, B. L.

1972 "Preliminary report of Taichung IUD4 Study." University of Michigan, Population Studies Center, mimeo.

Potter, R. G.

1966 "Application of life table techniques to measurement of contraceptive effectiveness," Demography 3: 297-304.

1967 "The multiple decrement life table as an approach to the measurement of use effectiveness and demographic effectiveness of contraception." Contributed papers, Sydney Conference, International Union for the Scientific Study of Population, 869-883.

Potter, R. G., *et al.*

1962 "Improvement of contraception during the course of marriage." Population Studies 15 (November): 160-174.

Ross, John A., Frederick F. Stephan, and Walter B. Watson

1969 A Handbook for Service Statistics in Family Planning Programs. New York: The Population Council, Inc.

Tietze, C.

1962 "Use effectiveness of contraceptive methods" in C. V. Kiser (ed.), Research in Family Planning. Princeton: Princeton University Press.

Tietze, C.

1967 "Intra-uterine contraception: recommended procedures for data analysis." Studies in Family Planning 18, Supplement (April): 1-6.

Tietze, C. and S. Lewit

1968 "Statistical evaluation of contraceptive methods: use effectiveness and extended use effectiveness." Demography 5: 931-940.

United Nations

1969a Report of the Expert Group on Assessment of Acceptance and Use Effectiveness of Family Planning Methods.
Asian Population Studies Series No. 4. United Nations Publication. Sales No. E.69.II.F.15.

1969b Evaluation of Family Planning Programmes. Asian Population Studies Series No. 5, chap. V, pp. 11, 12, esp. 56-63. E/CN.11/936.

ACKNOWLEDGEMENTS

Gratefully acknowledged is financial support from the Ford Foundation and The Population Council and the programming assistance of Mrs. Josephine De Hart.

Chapter 6
BIRTHS AVERTED

by
David Wolfers

I. INTRODUCTION

The term "births averted" has been employed by different authors to describe several different approaches to the measurement of the demographic impact of contraceptive programs. This chapter is confined to the relatively narrow field of estimation of the births averted from cohorts of acceptors of birth control by a single "segment" of use, defined as follows:

> "The estimation of births averted is a measure of the quantitative change in the expectations, assessed in terms of probabilities, of future births to a cohort of women resulting from the adoption or modification of birth control practice by them or their husbands."

It is important to distinguish these estimates from others which compare the future fertility of acceptors with that of non-acceptors. In these the contraceptive strategy of acceptors *after* termination of a segment of use is incorporated into the calculations. While estimates of births averted as defined may play a part in predicting or accounting for changes in fertility indices over time (period approach) or in establishing the temporal fluctuations caused by isolated bursts of contraceptive activity, they are not specifically addressed to these problems.

This method of quantifying the demographic significance of contraceptive use has proved useful in exercises to determine cost/benefit of family planning programs which employ the economic cost (or benefit) of marginal births as a measure of the value of demographic change. Estimates of births averted also provide convenient indices for the comparison of different program strategies, different

birth control methods and different age-mixes of acceptors. In addition, they have a place in the process of "target-setting" (See Chapter 2 and Chandrasekaran *et al.*, 1971) and in providing a necessary demonstration of the way in which program performance leads to demographic changes.

While brief reference is made to the calculation of births averted by abortion and sterilization, the problem chiefly considered is births averted by the use of contraceptive methods for which continuation is suited to analysis by life-table methods: mainly the IUD, but also oral contraceptives. As pointed out by Chandrasekaran and Karkal (1972), there are serious difficulties in handling the use of conventional contraceptives in this way.

The principal published approaches to calculating births averted (as defined) are described and discussed in detail in section IV and an attempt to reconcile them is made in section V. Earlier sections (II and III) examine the materials used in constructing these model solutions, first from the theoretical point of view of their justification and the considerations which dictate the ways in which they must be handled; second, in relation to the types and sources of data required and to the analytic or synthetic procedures applied to the data.

Among the important and to some extent controversial factors which must be brought together to construct a realistic model is the potential fertility of acceptors against which their reduced fertility after acceptance must be weighed. Its estimation requires the introduction of a considerable number of parameters, including the age-specific fertility performance of the population at the time of measurement, ways in which acceptors may be expected to differ in their fertility potential from otherwise similar groups of women, the extent to which acceptance is merely the substitution of a program for a non-program method, the incidence and prevalence of sterility amongst acceptors and, where possible, expected spontaneous change in fertility levels not attributable to programs. Pregnancy rates during contraceptive use, live-birth ratios, overlapping between contraceptive use and post-partum anovularity, duration of use of contraceptives, mortality, widowhood and divorce are other factors to be considered.

Where simulation models are used for estimating births averted, the substitution of fecundability estimates for potential fertility simplifies many of the problems involved, but the derivation of these estimates raises its own problems.

II. CONCEPTS

THE UNIT OF CONTRACEPTIVE USE

The unit of contraceptive use is one month's contraception by one couple.

A SEGMENT OF CONTRACEPTIVE USE

A segment of contraceptive use is an uninterrupted sequence of months of use of the same method by the same couple. Segments commence at adoption and terminate at abandonment, expulsion (for IUDs), exhaustion of supplies, pregnancy or death, whichever is soonest.

THE DEMOGRAPHIC EFFECT OF A UNIT OF CONTRACEPTIVE USE

The demographic effect of a unit of contraceptive use is to reduce, usually substantially, the probability of conception occurring

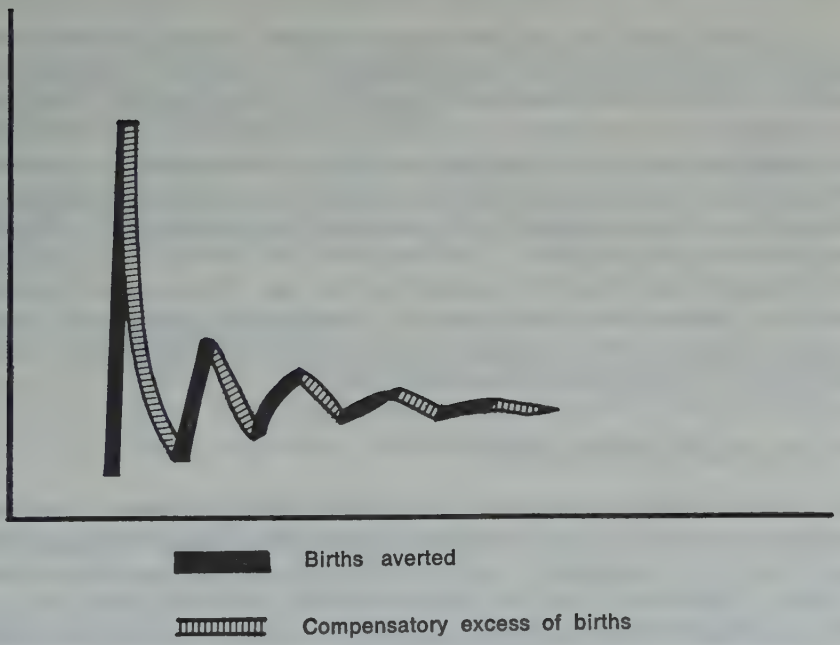


Fig. 6.1. *Timing of births averted by one unit (month) of contraceptive use.*

during that month and hence of a birth taking place seven to ten (usually nine) months later and to alter, in progressively diminishing degree, the probabilities of conception in all subsequent months of reproductive life. The birth-averting (or birth-preventing) value of a unit of contraceptive use is the arithmetic sum of all these changes in probability. Some of the changes are positive, representing an increased probability of conception in that month, others are negative. The sum, except in the case of a less efficient contraceptive being substituted for a more efficient one, will always be either zero or negative, confirming that contraceptive use operates to reduce the total probability of births, i.e., it averts births. Figure 6.1 shows these changes in monthly probabilities for a unit of contraceptive use of a one hundred per cent effective contraceptive by a woman in the fecundable state during the month of use.

BIRTHS AVERTED BY A SEGMENT OF USE

The birth-averting effect of a segment of use is the arithmetic sum of the effects of each month of use comprising the segment.

A "REPRODUCTIVE UNIT" OF TIME

A reproductive unit of time is the average interval between successive recurrences of the same stage of a woman's reproductive cycle, i.e., the interval between conception and conception, delivery and delivery or resumption of post-partum fecundity between one birth and the next. The reproductive unit is applicable only to women who retain the capacity to reproduce (i.e. are non-sterile) throughout its duration. It corresponds to the expected prospective birth interval and is strongly age-specific. The length of the reproductive unit may be simulated for groups of women by the use of the monthly probability of conception (fecundability), post-partum anovularity, pregnancy duration, foetal wastage and monthly probabilities of conversion from a non-sterile to a sterile state.

THE "DURATION PER BIRTH"

The duration per birth is only subtly distinguishable from the reproductive unit of time. This period includes in its calculation the contribution of women who, while non-sterile at the beginning of an interval, nevertheless do not conceive again before becoming sterile. It is equivalent therefore to the average contribution from their total reproductive lives which is required from women for each pregnancy.

The period is always longer than the reproductive unit of time and the prospective birth interval, and towards the end of the reproductive life it is very much longer.

RELATION BETWEEN DURATION PER BIRTH, SEGMENT OF USE AND BIRTHS AVERTED

It can be shown that the use of effective contraception commencing in the fecundable state for a segment of use equal to a woman's duration per birth averts exactly one birth from the total expected by her, and its use for a given proportion of the duration per birth has that proportionate probability of averting a birth. The computation of births averted from a cohort of women by a number of segments of contraceptive use therefore consists of the expression of this use in terms of corresponding durations of births which it comprises, with appropriate corrections for inadequacy of the contraceptive and periods of use by women not in the fecundable state. The most important theoretical problem in the estimation of births averted is the accurate computation of appropriate durations per birth for cohorts of acceptors.

CHARACTERISTICS OF METHOD AND USERS

The quantitative value of the probability changes induced by contraceptive use is importantly influenced by a number of characteristics of the method and the client.

1. *THE EFFECTIVENESS OF THE METHOD*, as measured by proportionate reduction in probability of pregnancy in one month of use by a woman exposed to the risk of pregnancy, is usually a function of:

a) the method itself;
b) the length of time it has been used by the individual client;
c) age and parity of the client (Zatuchni, 1970: 297 and 329);
d) physical and social characteristics of the client.

Certain methods (abortion and sterilization), however, are virtually immune to the influence of *b*), *c*) and *d*).

2. *THE PHYSIOLOGICAL STATE OF THE WOMAN DURING THE MONTH OF USE*. A woman of reproductive age may at any time be in any one of the following states:

(i) Fecundable - Having a greater than zero probability of conception, dependent on frequency of intercourse, quality of mate's sperm, timing of intercourse, chance factors, etc.

(ii) Temporarily infecundable - With no probability of conception for reasons which are spontaneously reversible, for example, anovulatory cycle, absence of spouse.

(iii) In a state of post-pregnancy anovulatory immunity [an important subgroup of (ii)].

(iv) Temporarily sterile - With no probability of conception, for pathological, but reversible, reasons, or husband sterile.

(v) Permanently sterile - With no probability of conception, either for non-reversible pathological reasons, for example surgical intervention, or because of termination of the capacity to conceive as a result of the passage of time.

(vi) Pregnant.

The use of contraception by a woman in any but the first state will have zero effect on the probability of future births or conceptions.

The problem of determining the state of any individual woman during any individual month of contraceptive use is not susceptible to practical solution. For cohorts of women, however, estimates of the numbers at any one time in states (iii), (v) and (vi) can be derived, while the proportions in any one month in states (i), (ii) and (iv) can be subsumed within general estimates of conception probabilities.

Married women, women who have never been delivered of a live birth, women at the time of acceptance of contraception and users of contraceptives form four identifiable groups amongst whom the distribution of these six states is different.

3. *THE AGE OF A WOMAN* is strongly correlated with the monthly probability of her having a conception leading to a live birth. A number of mechanisms are involved:

a) There appears to be an age-related fecundity curve—defined as the probability of conception under standard conditions—which rises steeply to a maximum from menarche to late teens, flattens in the early twenties and declines continuously thereafter.

b) As a general rule, frequency of intercourse declines progressively with increasing age, particularly for men. In stable unions this decline will, of course, be parallel for women. The effect of reduced frequency of intercourse within limits is to reduce fecundability or the real monthly probability of conception.

c) The mean duration of post-partum infecundity for groups of women shows a progressive increase with age (Potter, 1963).

d) Foetal wastage rates increase with age of the expectant mother.

e) The incidence of pathological permanent sterility and presumed physiological permanent sterility are cumulative and accelerative so that at each age a larger proportion of women is in those states than at younger ages, and during each month of consecutive contraceptive use, higher proportions of older women become permanently sterile than of younger women.

f) The proportion of women already using contraception or otherwise ensuring or attempting to ensure that they will bear children only at will is found to vary with age.

g) The effectiveness of contraceptive methods, whether dependent on or independent of the diligence with which they are used, has been found to vary with age [see (1) above].

4. *SUBSTITUTION.* This term is used to describe the situation where the adoption of contraception in the program context replaces the use of other methods of birth control outside the program. What we might term a second derivative of substitution is the adoption of contraception inside the program by people who would have adopted contraception outside the program if the program had not existed.

Neither of these two forms of substitution can be accurately measured. At any given time a population of acceptors can be divided, with reasonable accuracy, into three categories—past users, present users and never users. A uniform finding of studies of family planning acceptors, against the background of the population of married women of reproductive age, is that previous contraceptive use predisposes to acceptance, although a large part of this effect is attributable to the generally higher parity of previous contraceptive users and the strong positive correlation between parity and acceptance of program contraceptives (Wolfers, 1970a: 62). To establish that a woman was using a non-program method of contraception up to the time of acceptance, or, still less, at some time before acceptance, is not, however, to establish that she would have continued to use it either for any time, or for as long a time as she uses the program method, had she not become an acceptor. The proportion of ever-users who are also current users of contraception has been found to vary greatly, both between and within countries. Thus amongst Ecuadorian women interviewed in 1971, S. Scrimshaw (1973: 175) found that past experience of contraception was reported by 44.5 per cent of the female population and current experience by only 26.5 per

cent; for men the respective figures were 62.5 and 26.4 per cent. The West Malaysian Family Survey (National Family Planning Board, 1968) revealed that in rural areas only half of the 6 per cent of women who had ever used contraception were current users. In the towns, where practice was much more common (40 per cent for large towns and 28 per cent for small ones), current users amounted to 75 per cent of ever-users. In Singapore, 72 per cent of acceptors in 1963 were still (or again) using contraceptives in 1967, 88.5 per cent of them obtaining supplies from the sources provided by the government program which began in 1966 (D. and H. Wolfers, 1968). On the other hand, as Freedman and Takeshita (1969: 304) have demonstrated for Taiwan and Wolfers (1970a: 111) for Singapore, discontinuers from within a program are very likely to continue to use birth control, either by different methods within the program, or by methods from outside it.

These figures demonstrate that the past or current use of contraception is no reliable guide to future practice and that the appearance of substitution may be very misleading. Before the introduction of oral contraceptives and IUDs, the discontinuation rates of conventional contraceptives, particularly in developing countries, were very high indeed. Mean periods of use of diaphragm, foam and condom respectively in the Singapore clinic referred to were 4.0, 6.5 and 9.5 months.

It is therefore almost a safer hypothesis that when a current contraceptor changes to a program method, she was on the point of abandoning the old one than that she would have continued to use it. The quantification of this and the more conditional second derivative of substitution remains an effort of judgement rather than a result of factual knowledge. Over-elaborate attempts at quantifying substitution effects serve only to lend spurious verisimilitude to births-averted calculations without adding to their realism.

5. *SOCIAL VARIABLES.* Although they take effect through intermediate variables, the influence of certain social variables on the birth-preventing value of contraceptive use is sufficiently consistent and important to warrant their consideration as a separate phenomenon. The significant variables include ethnic group, education and social class. In different categories under these headings are likely to be found differences in:

a) effectiveness of contraceptive practice;
b) persistence or continuation of use;
c) previous contraceptive practice and hence weight of substitution factors;

d) age-distribution of acceptors;
e) baseline fertility values.

POTENTIAL FERTILITY

As the unit of contraceptive use has the dimension of time—one month—it is helpful in non-simulation exercises to express fertility parameters in the same dimension. While the normal fertility measures are calculated as births per unit of time, it is conceptually more correct for calculations of births averted that these be inverted to become time per birth. To achieve this, Potter (1969a) has devised a measure termed "duration per birth" (see above), Lee and Isbister have accepted a direct inverse relationship between age-specific marital fertility rates (ASMFRs) and birth intervals, while Wolfers (1969) has worked directly with observed birth interval data.

The actual relationship between fertility measures expressed in births per unit of time and units of time per birth is far more complex than at first appears. Louis Henry (1961) appears to have made the first statement on the subject:

> "Le taux de fécondité des couples fertiles est à peu près égal à l'inverse de la moyenne des intervalles dont le début ou la fin ou les deux tombent dans le groupe d'âges considéré. . . ."
> [The fertility rate of fecund couples is as a close approximation equal to the reciprocal of the mean of those (live-birth) intervals of which the beginning, the end or both fall in the age-group considered. . . .]

This statement has been occasionally misread (e.g. Lee and Isbister, 1966: 746) as implying that the average birth interval "is given by the inverse of the fertility rate", while in reality it does not reveal any usable relationship between the two quantities except under unrealistic conditions (Potter, 1970). Henry speaks of the "fertility rate of fecund couples" which is a non-measurable abstract concept. Fertility rates can be calculated for all sorts of selected groups, but not for fecund couples specifically, for these cannot be identified. In addition, the group of birth intervals described, commencing or terminating within a certain age-group, does not correspond to the measure of future potential fertility required for birth-averted calculations where only future (or prospective) intervals, i.e., those commencing at the specified age, are relevant.

The baseline fertility characteristic, duration per birth, will be taken or derived from some known fertility parameter(s) of the population to which the cohort belongs. The selection of baseline potential fertility is, of course, crucial to the realism of the quantitative

results of a model, but the controversy over how this should be done is far from being resolved. Among the problems besetting this area are:

(1) *GENERAL*

a) Differential fecundity. As a generalization, women adopting contraceptives have more children for age and duration of union than all women with otherwise similar characteristics. This suggests, but does not establish, that they have a higher than average potential fertility.

b) Substitution. Women who have previously used contraceptive methods, abortion or abstinence are over-represented amongst acceptors of contraceptive programs. These women can properly be assumed to have a lower than average potential fertility. Likewise, acceptors may be assumed to include a disproportionately large number of women shortly due to engage in birth-limiting practices, even without the program's facilities.

c) Trend. In rapidly moving social panoramas, spontaneous changes in fertility behaviour are common. Attempts to give validity to assessments of the achievements of a program involve making allowance for pre-evident trends in fertility values. It cannot be assumed, without running the risk of performing a mere abstract mathematical exercise, that baseline fertility values derived for an initial year will remain applicable for future years.

(2) *FERTILITY RATES*

a) *Alignment.* At the time of adopting contraception, women are not mixed in physiological states in the same way as the general population who form the denominator for the ASMFRs, but are highly selected for being either fecundable or in a state of post-partum anovularity. The proportion of contraceptors who are permanently sterile and those who are pregnant will be very much smaller than in the population at large.

b) *Parity.* Nulliparous women do not form part of contracepting populations in developing countries, but the contribution of first births to ASMFRs at early ages is quantitatively important. The mathematical relationship between age-specific marital fertility rates and "duration per birth" is modified in a major way by the proportions of first births contributing to the fertility rates (Wolfers, 1973).

(3) *BIRTH INTERVALS*

Truncation effect. Average prospective birth intervals become shorter at older ages because potentially longer intervals are not

realizable due to the intervention of permanent sterility. The use of birth intervals therefore exaggerates potential fertility in older age groups.

(4) *COMPUTER SIMULATION MODELS*

Derivation. Where fecundability values are employed to measure baseline fertility, these are, by nature of the quantity, secondary deductions from other fertility data. They are derived by fitting to the known fertility schedules of a population sets of parameters for fecundability, post-partum anovularity, sterility, widowhood etc., which when fed into a computer model will closely duplicate the fertility schedules. For any fertility schedule, there is an unlimited number of solutions, so that the fecundability estimates are in some degree arbitrary. Where hard data can be derived for some parameters —post-partum anovularity or foetal wastage for example—the range of realistic values available for fecundability shrinks and more confidence can be placed in the deductions from such systems (see Venkatacharya, 1971b). The questions of how much heterogeneity of fecundability and anovularity to introduce into these models are also important, but require arbitrary solutions.

STERILITY

Our knowledge of the way in which permanent sterility intervenes to terminate the reproductive life is very incomplete. Only three items of secure knowledge exist:

a) a small proportion of women in all societies are primarily sterile—i.e., do not, despite active sexual lives, produce children;

b) there is a risk of the development of sterility at every age of child-bearing life and, because of the permanence of sterility, the proportion sterile is cumulative with age;

c) the risk of development of sterility increases with age, so that the rate of change of proportions sterile accelerates throughout the child-bearing ages, becoming steep at ages over 40 and reaching 100 per cent sterile at about age 52.

Questions which have still to be answered include:

(i) What are the parameters of the incidence and prevalence of sterility in specific communities and what are the factors (for example, ethnic, geographic/climatic, dietary, demographic, social, morbid) which determine them?

(ii) Does childbirth itself engender a significant additional risk of sterility?

(iii) Does pregnancy provide a period of immunity from the risk of sterility?

(iv) Does induced abortion heighten the risk of the development of sterility and if so, how much of this is dependent on the way the abortion is performed?

The direct measurement of sterility is not at present possible, there being no diagnostic criteria to distinguish the fecund from the sterile in the absence of specific pathological states.

Barrett (1971) has shown that truncation effects are present in the case of sterility measurements whether derived by analysis of proportions of marriages childless by age at marriage, or by analysis of age at last child birth. The effect of these is always to overestimate the proportions sterile.

The importance of sterility to calculations of births averted lies in the self-evident fact that contraceptive use by sterile women has no birth-averting effect. Women who believe they are sterile do not adopt contraception, but this is by no means the same proposition as saying that women who are sterile do not adopt contraception. Women who adopt contraception usually lose the capacity to form an opinion of whether or not they have become sterile.

TIMING

One problem in the analysis of births averted which has either been neglected or inadequately treated in early studies is the time element in averted births. In the models of Potter and Wolfers, while a net value has been placed on the births averted, no indication is available of the fluctuations in numbers of births to be expected in a community as a function of elapsed time. This is indeed a thorny problem. 1,000 units of use of a 100 per cent effective contraceptive in 1,000 fecund women for the month of January will prevent a number of births (say between 100 and 200) which would have taken place in or around October. If contraceptive use stops there, secondary effects will be that those women who would have given birth in October will now have the possibility, generated by their contraceptive use, of giving birth at a later date. The 100-200 missing births in October will be compensated for by approximately the same number of additional births spread over the ensuing months. The deferment of these births will lead as a natural consequence to the deferment of the next births to the women concerned and hence to a reduction of births at a later period, followed again by a compensatory excess (see Figure 6.1). While neat theoretical probability models can be

used to demonstrate that this oscillation will persist to the end of the reproductive lives of the women involved, in practice the multiple irregularities ensure that these damp down within at most two or three birth intervals (Barrett, 1972).

While the task of determining the timing of averted births attributable to a single unit of contraceptive practice, when solved, does not yield useful information, the more realistic problem of timing births prevented from aggregates forming a continuum of contraceptive use is at once of more practical value and easier to analyse.

K. Venkatacharya (1971a, 1972) has devised matrices indicating the annual probabilities of births to women coming under observation in susceptible (fecundable) states at different ages, which can be derived from age-specific marital fertility rates. These tables can be used to establish the timing of births prevented by sterilizations and, with a little more labor, those averted by contraceptive use of limited duration. Although the actual model used by Venkatacharya employs a number of simplifying assumptions which detract from its value as a tool for quantifying births averted, the distribution of these over time appears to emerge with reasonable precision.

An alternative solution to the timing of births averted lies in direct computer simulation by means of Monte Carlo type models from which the annual, or even the monthly, birth fluctuations induced by any specified birth control strategy can be obtained. This suggestion, together with an explicit consideration of problems of timing, is found in Chandrasekaran *et al.* (1971).

MEAN PERIOD OF USE

Data may be handled in one of two ways, either by the computation of a single value—the mean period of use (perhaps using several different sub-groups of acceptors)—or by making month-by-month calculations employing continuers in each month as the population of contraceptive users.

The former method, while mathematically neater, introduces an additional approximation into the calculation, the mean fertility of users being taken as the fertility of women at the mean age during use.([1]) It also carries a heavy temptation to assume regular decay-

([1]) The approximation will be valid provided either that:

(i) The change in fertility potential with age for the age-range of users is trivial (which will apply if mean duration of use is relatively short and range small, or if the middle age-range of reproduction only is involved).

or

(ii) The change in fertility potential with age follows a parallel path to the decay curve of contraceptive use (e.g. if both are exponential).

curves of continuation which may or may not correspond closely to reality (see section III.6).

In the case of sterilizations, the computation of mean periods of use (from operation to death, widowhood, separation or sterility, whichever is soonest) to derive mean age during use (for specific age-groups at operation) and hence a mean fertility, would lead to serious inaccuracy.

LONG-TERM AND INDIRECT EFFECTS

Some birth-averted calculations, used as justification for proposed programs, project births averted by a suggested continuing strategy into the long-term future—20 or 30 years. We would suggest that no trace of reality clings to such projections for the dimensions of substitution and trend in demographic indices must get entirely out of hand when looked at with such a perspective. This problem also casts a shadow over summary calculations of births averted by sterilization.

SECOND-GENERATION EFFECTS. The prevention of births which would otherwise have taken place reduces the number of potential parents in the next generation, and hence may be presumed to generate additional births averted some 20 years later. Indeed, carried to its illogical extreme, provided the net reproduction rate is greater than 1, any birth prevented generates an infinity of births prevented into the future. Happily not much effort has been spent on the numerical assessment of second and later generation effects, for this is not a field that lends itself to profitable exploitation. It is necessary in making such calculations to assume parameters for the fertility of the population 20 and more years into the future and, as the wreck of innumerable attempts at demographic prediction show, it is not possible to do this with any useful degree of accuracy.

Potter (1972a) has done some work on intergenerational effects, but not with any view to practical evaluation of programs.

III. DATA: METHODS OF HANDLING, SOURCES AND LIMITATIONS

FERTILITY

The problem of determining the fertility expectation of a group of acceptors, and expressing it as duration per birth, still has no agreed solution. Potter (1969a) has summed up the first part of the problem: "acceptors ... are doubly selected, having higher

fecundity than average for their age class and partly for that reason also higher than average interest and initiative with respect to family planning". The second point relates to the unquantifiable second derivative of substitution [see above, section II.8(4)], while the first is very controversial.

Numerous writers, for example Ross (1965), Potter *et al.* (1968), Rele and Patankar (1969) and Jain (1969), have found that family planning acceptors have higher than national fertility rates or parity for age schedules. This is a fully expected finding, given that micro-demographic pressures play a prominent part in motivation to acceptance. It does not, however, follow that a group identified *post facto* as having higher than average past fertility also has higher than average current fertility nor if it has, is it clear to what extent its higher fertility is a result of higher fecundity and to what extent the result of non-systematic variability. As Ridley *et al.* (1969) have demonstrated with their computer simulation model, "Repsim", any plausible model of acceptance behaviour will lead to the development of substantial fertility differences between acceptors and non-acceptors even when they are given identical values for all probability parameters of reproduction. Brass (1970) has demonstrated in computer simulations with heterogeneous fecundability that the chance element heavily outweighs the systematic, especially for short marriage duration.

A large measure of the genuinely higher fertility expectation of acceptors derives from the exclusion from their ranks of women who assume themselves to be sterile. The longer the elapsed time since the last childbirth, the lower the probability that a woman will accept contraception. In the pioneer study of births averted, Lee and Isbister (1966) assumed a 20 per cent higher fertility of acceptors than for the general population, basing their figure on observations made in Taichung. The actual differential (between recent past fertility of acceptors and population marital fertility rates) was highly age-specific, ranging from 15 per cent at ages 25-29 to 64 per cent at ages 35-39 (Ross, 1965). At yet higher ages, as sterility becomes a majority phenomenon, very high values are reached. Later figures from Taiwan show differences soaring to 275 per cent at ages 40-44 (Chow, 1968). These differentials are very different at different ages, and their significance also changes with age. At young ages sterility contributes a minor part of the difference, its bulk flowing from differential fecundity and random variation. At high ages, sterility becomes overwhelmingly the most important factor. It is not feasible to disregard such differences, nor to ignore their age-specificity. One way of coping with the situation is that employed by Potter (1969a) which employs the reciprocal of the recent past age-specific fertility

rates of acceptors only as the basis of computation. Not only is this method heavily demanding in terms of accurate data (readily available in Taiwan, but not in many other places) but it incorporates a bias towards the inclusion of women with high fertility performance as a random expression of normal fecundity leading to overestimation of potential fertility. Another way of separating fertile from infertile women and including only the experience of the former is Wolfers' (1969) use of birth intervals as a measure of future performance, also heavily data-dependent and subject to the bias introduced by truncation of birth-intervals at later ages.

Methods for elimination of these biases have been proposed recently by Wolfers (1973). The first of these involves the processing of age-specific marital fertility rates (from total populations) by a formula which is designed to eliminate the contribution of sterile women from the rate and to compensate for the proportion of first births:

$$D_j = \frac{36{,}000\,(1-S_j)}{B_{j+2}\,(3-N_{j+2})} \qquad (1)$$

where

D_j is duration per birth, in months, for women aged j years;
S_j is the proportion of women aged j who are sterile;
B_j is the age-specific marital fertility rate for age j;
N_j is the proportion of all births to women aged j which are first births.

The results of the application of this formula, which also corrects for the contribution to marital fertility rates made by first births and incorporates an automatic adjustment for the development of sterility during use, is to bring potential fertility estimates generally slightly *below* age-specific marital fertility rates up to age 30 or 35, and above marital fertility rates at more advanced ages.

The second methodological innovation is the use of a multiplier matrix to apply to birth-interval frequencies in order to eliminate truncation bias. The opportunity to examine the truncation correction matrix by applying it to data has not yet arisen.

The most common measure of potential fertility employed in calculations of births averted is the direct use of the reciprocal of age-specific marital fertility rates, either unmodified or subject to more or less arbitrary multiplication by a factor recognizing the higher fecundity of acceptors (Talwar *et al.*, 1969). This approach certainly has the advantage of simplicity.

Fundamental data from which the calculation of potential fertility can be made derive either from available vital statistics—ASMFRs or

ASFRs matched with proportions married—or from special studies designed to provide detailed fertility data.

1. *RATES.* The majority of countries do not produce ASMFRs on an annual basis—the vital registration data required for this being comprehensive and relatively complex. Needed annually are:

a) births by age of mother and legitimacy,
b) marriages and divorces by age of bride,
c) deaths by age, sex and marital status,

together with firm census data on age, sex and marital status and reasonable migration statistics.

Estimates are commonly made by dividing age-specific fertility rates by proportions of women married in each age group obtained from recent census data. Freedman and Adlakha (1968), using hospital records to determine the age distribution of women bearing children, have derived sets of ASFR and ASMFRs for Hong Kong (1965 and 1966) where birth registration is deficient in not recording age of mother.

Fertility rates are subject to all the inaccuracies which can affect enumeration and age-reporting in census and registers. Marital rates are characteristically less accurate than general rates, particularly in countries where formal marriage is uncommon. In most Latin American and Caribbean countries, marital fertility rates are seriously misleading. Rates derived from special registration areas are often very far from representative of whole countries.

2. *SURVEYS.* Special surveys require very meticulous sampling procedures in order to ensure that their results are representative, but on the whole provide the most reliable indices of fertility in developing countries. Retrospective fertility surveys show consistent decay in completeness the longer the period concerned. Surveys to establish contemporary age-specific fertility rates need not, however, probe further than one year into the past.

The investigation of the fertility rates of groups of acceptors isolated from the general population is dogged by artifacts related to the shorter open intervals of acceptors. Furthermore, a substantial but unknown proportion of the differences found between past fertility of acceptors and the population at large is of a chance nature and has no prognostic value.

3. *MATCHING STUDIES*, of which that of Chang *et al.* (1969) is almost the only example (the matching study of Haynes *et al.* (1969) was retrospective and used in default of adequate data on past fertility

of acceptors; it is therefore no more than a superior example of borrowed data.) Follow the *subsequent* fertility of a group of non-acceptors after matching for age, parity, open interval and education of acceptors. In the circumstances of such a study, in which none of the matches is permitted to be pregnant on recruitment to the group, the future fertility of the matches will be oscillatory and not reach an asymptotic level for several years. It does not, therefore, provide a set of rates which can be used in "conventional" births-averted calculations.

If, as is done by Chang *et al.*, births averted are calculated by direct comparison of post-acceptance fertility of acceptors and matches, without regard to whether or not they continue or discontinue IUD use, become sterilized, have abortions etc., we find ourselves no longer studying births averted by use of a contraceptive method but instead, the differential fertility or two groups in a population distinguished by the decision made or not made at a particular time, that is, the significance as a predictor of future fertility of the decision to accept or reject the offer of an IUD.

The availability of data for matching studies of this kind depends on the pre-existence of very comprehensive and detailed registration data, such as are found in Taiwan and hardly elsewhere.

4. *BIRTH-INTERVAL DATA* are rarely found fortuitously. Increasingly, however, fertility surveys are being conducted in such a way that birth-interval data can be extracted from them, although memory bias may require that their results be accepted with caution.

The interval summaries required for the calculation of births averted are age-specific mean prospective intervals (births). Prospective intervals are those commencing at the given age and are not even approximately interchangeable with retrospective intervals (terminating at that age). Mean intervals for births are distinguished from mean intervals for women as the means of birth-interval distributions in elapsed time, i.e., the means of all intervals terminating between set dates (collected most readily from birth records or maternity hospital data). Mean intervals for women are collected in such a way that all women in a sample contribute equal numbers of intervals, for example, by recording the last birth-interval of all women surveyed. Mean intervals for births are always shorter than those for women. The relationship between the two is discussed in Wolfers (1968).

It is important to realize that for deriving fecundability estimates from birth intervals, mean intervals for women are required.

5. *FECUNDABILITY ESTIMATES.* In computer simulation mod-

els, fecundability estimates are required for sub-groups. It is perhaps easier to describe how such estimates cannot be made than how they can. Fertility rates, marital or otherwise, require extensive processing to exclude the components introduced by sterile women, contracepting women, women not exposed to the risk of pregnancy, post-partum amenorrhea (but not for nulliparous women), pregnancy duration, foetal loss, etc. In practice the task is too speculative to provide useful results. To use birth intervals more must be known of the distribution of post-partum amenorrhea and the time components of foetal loss than is usually available.

The direct and rigorous study of arithmetic mean fecundability is achieved by following groups of women, all ovulating, from the commencement of exposure to pregnancy risk until all (who are going to) become pregnant. This requires either virgin marriage cohorts or women who abandon contraception in order to conceive (Vincent, 1965; Westoff *et al.*, 1961). Such studies are meticulous research procedures and not suitable for field evaluation.

In practice, fecundability has been studied by observing the proportions of groups of women presumed to be ovulating who become pregnant in successive months over relatively short periods. Consistently, smaller proportions of non-pregnant remainders become pregnant with each succeeding month, a manifestation of the heterogeneity of fecundability in all populations.

By appropriate mathematical analysis, such data can be used to derive estimates of fecundability distributions, as well as mean fecundabilities. Mukerji and Venkatacharya (1967) have pointed out that true biological fecundability is not necessarily the most appropriate parameter for computer simulation and that a conception probability may be derived from ASMFRs "such that, when used with a set of other relevant input parameters, the simulated set should yield age-specific marital fertility rates as close to an assumed set as possible" and they have developed a method for doing this.

In general, neither denominators in the form of numbers married by age, nor numerators as births by age of mother, nor yet intervals between births by age, are accurately known. For a long time to come, births averted calculations will rely on crude estimates, sample surveys, or borrowed data. These deficiencies are, if anything, magnified when fecundability estimates are required, as for computer estimates.

POST-PARTUM AMENORRHEA

Particularly in post-partum contraceptive programs, adoption of contraception or sterilization frequently occurs while a woman is still

Table 6.1. *Computation of overlap between contraceptive use and post-partum anovularity*

Open interval	*Proportion adopting*	*Proportion ovulating*	*Proportion ovulating by month after adoption*											
			1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
i	Y_i	Z_i	Y_iZ_i	Y_iZ_{i+1}	Y_iZ_{i+2}	Y_iZ_{i+3}	Y_iZ_{i+4}	Y_iZ_{i+5}	...	...	...	...	...	$Y_iZ_n = Y_i$
1	0.05	0	0	0.0025	0.005	0.0075	0.01	0.015	0.02	0.025	0.0325	0.04	0.045	0.05
2	0.05	0.05	0.0025	0.005	0.0075	0.01	0.015	0.02	0.025	0.0325	0.04	0.045	0.05	0.05
3	0.10	0.10	0.010	0.015	0.02	0.03	0.04	0.05	0.065	0.08	0.09	0.10	0.10	0.10
4	0.10	0.15	0.015	0.02	0.03	0.04	0.05	0.065	0.08	0.09	0.10	0.10	0.10	0.10
5	0.20	0.20	0.04	0.06	0.08	0.10	0.13	0.16	0.18	0.20	0.20	0.20	0.20	0.20
6	0.20	0.30	0.06	0.08	0.10	0.13	0.16	0.18	0.20	0.20	0.20	0.20	0.20	0.20
7	0.10	0.40	0.04	0.05	0.065	0.08	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10
8	0.10	0.50	0.05	0.065	0.08	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
9	0.05	0.65	0.0325	0.04	0.045	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
10	0.05	0.80	0.04	0.045	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
11	—	0.90												
12 (n)	$\sum_{i=n}^{i=\infty} Y_i$	1.00												
A Sum			0.2900	0.3825	0.4825	0.5875	0.695	0.790	0.870	0.9275	0.9625	0.985	0.995	1.00
Complements			0.7100	0.6175	0.5175	0.4125	0.3050	0.210	0.130	0.0725	0.0375	0.015	0.005	0.00
Sum of complements = mean overlap with anovulatory period = 3.0325 months														
Potter's method: Mean p.p.a. = 6.95 months. $0.2 \times 4.95 + 0.5 \times 1.95 = 1.965$ months														
Wolfers' method: Mean open interval to acceptance = 5.5 months. Mean overlap = 2.85 months														
Mean duration of p.p.a. only for those ovulating			4.162	4.51	5.04	5.50	5.92	6.26	6.53	6.72	6.82	6.89	6.93	6.95

in an anovulatory state following her last delivery. The period of contraceptive use by a cohort which overlaps anovularity must be deducted from the total contraceptive use before this can be made a basis for calculating births averted. It is self-evident that use during such periods of overlap does not contribute to averting births.

Three methods of computing the loss in useful contraceptive effect due to overlap between contraceptive use and post-partum anovularity have been described. The first two of these are seriously inadequate:

(1) *POTTER (1969a)* computes mean overlap as follows:

a) obtain mean duration of post-partum amenorrhea;
b) class acceptors in 3-month groups by interval between last birth and acceptance;
c) subtract mid-points of these classes from mean amenorrhea (ignoring groups with mid-points greater than mean amenorrhea);
d) multiply each of the resulting values by the proportion of acceptors in the corresponding class and the proportion still effectively wearing the device at the point of mean amenorrhea;
e) sum to give mean overlap.

The flaw in this method can best be illustrated by assuming that all acceptance takes place within a three-month interval group after delivery whose mid-point coincides with mean post-partum amenorrhea. Then no allowance will be made for overlap, even though 50 per cent (approximately) of adopters will be amenorrheic at the time of adoption. The unstated underlying assumption that negative overlap cancels out positive overlap is not true.

(2) *WOLFERS (1969)* computes overlap, month by month after acceptance, by assuming that all women become adopters at the mean delivery-acceptance interval and matching this with an assumed distribution of post-partum amenorrhea by month after delivery. The correction is applied only to surviving effective users for the month in question.

When there is considerable spread in delivery-acceptance intervals, this method also underestimates overlap, although less seriously than the former method.

(3) *WOLFERS (1971)* proposed a method which is accurate provided that one can assume independence between the two distributions and is feasible when data for both are available. If only a mean value for post-partum anovulation is available, a theoretical distribution may be fitted and the method applied to this. It is illustrated in Table 6.1 for hypothetical distributions of delivery-adoption intervals and durations of post-partum amenorrhea.

In order to correct the mean overlap for discontinuation, it is merely necessary to multiply each entry in the "complements" row by the proportion of original acceptors still usefully retaining the device by the corresponding month after adoption. The last row in the table provides values, if required, for correcting durations per birth (see section IV.3).

Very often, information on the distribution of durations of this period is arbitrarily estimated or borrowed from (presumed) similar populations. A number of longitudinal studies from different areas do exist, and the distribution has also been studied by serial cross-section. The duration of post-partum amenorrhea has been shown to be very sensitive to breast-feeding practice and consequently, in mixed populations, has a bi-modal element (Sharman, 1966: 97; Potter *et al.*, 1965).

It is also likely that parity, nutrition, disease and post-delivery coitus customs all exert influence.

It is held by some (see Sharman, 1966: 70-87) that a variable number (with a mode of 1) of menstrual cycles after post-partum resumption of menstruation are anovulatory. If this is correct, anovulatory months should, of course, be added to amenorrheic months in allowing for overlap with contraceptive use.

ACCIDENTAL PREGNANCY

In handling pregnancies occurring accidentally during contraceptive use, Potter (1969a) has clearly demonstrated the correct approach. This is to assign for each pregnancy a "penalty" or deduction of months of useful contraception, equivalent to the mean period of ovulatory exposure to pregnancy risk required to yield one pregnancy in the absence of program contraception. This is calculated directly from durations per birth by subtracting nine months per pregnancy, the mean period of post-partum amenorrhea, and an allowance for the contribution of foetal mortality to durations per birth [set at one month for ages 20-34, 2 months for ages 35-39 (2)]. This period of ovulatory exposure is then divided by the ratio of pregnancies to live births to avoid penalizing for pregnancies destined to end in abortion. Pregnancy rates are found by extrapolating from life-tables of contraceptive use-effectiveness.

(2) Direct measurement on an Ecuadorian population gave values increasing from 0.9 months (age 15-19) to 3.3 months (age 40-45) (Wolfers and Scrimshaw, 1974).

FOETAL LOSS

Foetal loss is known to increase progressively with maternal age. Estimates of the proportion of pregnancies not terminating in live births are frequently made by arbitrary modification of values derived from the few (usually Western) detailed studies. Retrospective survey data of this quantity taken from high-fertility, low-education communities are likely to be almost worthless. Even neglecting (as is usual) miscarriages in the first six weeks of pregnancy, longitudinal prospective studies are very major undertakings.

STERILITY

Sterility estimates enter repeatedly into calculations of births averted:

a) In designating baseline fertility such that the selected rates or durations per birth relate to groups with proportions sterile approximating to proportions of acceptors sterile, or alternatively relate to non-sterile groups only. This has been discussed above in section III.1.

b) In computing the proportion of acceptors who are sterile at the time of acceptance. This has been dealt with in the literature in two ways:

(i) Potter (1969a) computes acceptors sterile at acceptance on the basis of a monthly risk of sterility applied to the mean period between last delivery and acceptance (taking the monthly risk as 1/60th of the proportion becoming sterile within a five-year age-group).

(ii) Wolfers (1969) applies a monthly risk of sterility (taken as 1 minus the 60th root of the proportion not becoming sterile within a five-year age-group) to the mean period between last *conception* and acceptance.

The additional nine months of risk of sterility allotted by Wolfers lead to significant differences between the two approaches at later ages—affecting total births averted to the extent of 3 per cent at age 35-39, 8 per cent at 40-44 and 36 per cent at 45-49.

c) In allowing for the development of sterility among acceptors during the currency of use. This has also been considered by both Potter (1969a) and Wolfers (1969), each of whom has applied a monthly conditional probability of sterility to achieve attrition of users. While Wolfers uses the same probabilities as for the computation of proportions sterile at commencement of use, Potter, whose

method requires exponential expression (3) of competing risks of "end of useful retention of device" (EUROD), derives the probability as follows:

s, the monthly conditional probability of sterility is found from

$$S_{x+2.5,x+7.5} = 1 - e^{-60s} \qquad (2)$$

where

$S_{x+2.5,\ x+7.5}$ is the risk of developing sterility between ages $x+2.5$ and $x+7.5$, and applies to insertions at ages between x and $x+5$.

For a discussion of the merits of including the nine months of pregnancy in the development of sterility before adoption, see section V.

Understanding the nature and incidence of sterility continues to evade us. While it is easy to define sterility as zero fecundability persisting for the remainder of life, this definition does not lend itself to operational measurement. Operational definitions related to actual child-bearing performance leave difficult problems of adaptation to calculations and models.

Data for the proportion of couples sterile by age of the wife are almost universally lacking. Agarwala has computed data for India and Louis Henry for a group of European populations. Two methods have been described, employing

(i) the proportion of couples completing their reproductive lives childless by age at marriage,
(ii) the age at last childbirth for non-contracepting couples (Vincent, 1950; Henry, 1953a, b).

Barrett (1971) has demonstrated how, whatever operational measure of proportions sterile is employed, simple analysis of observed proportions sterile will lead to exaggerated estimates as proportions sterile rise with age. The best hope of achieving some accuracy in this area is by the use of computer simulation models to quantify the inevitable discrepancies. With sterility occurring either menopausally or pre-menopausally, an unknown proportion being attributable to sterility of the male partner and the contribution of disease as a cause little explored, the hope of finding a simple universally applicable formula is slim.

(3) Exponential decay curves employed in quantifying sterility and contraceptive continuation are convenient mathematical fictions to describe events which, epistemologically, are manifest in discrete monthly rates. The geometric series is conceptually more correct.

CONTRACEPTIVE USE

In this category of information is required:

a) Continuation rates by month after adoption,
b) Pregnancy rates by month after adoption.

While of interest, the further breakdown of discontinuation rates by, for example, removal, expulsion, medical/non-medical reason, to have another child, end of need for contraception, etc., has limited relevance to birth-averted calculations.

Rates of this kind cannot be obtained from service statistics derived from clinic records as the experiences of those who report back to clinics differ greatly from the experiences of those who do not. This applies in different ways to all forms of contraception. Sample surveys of acceptors at varying times after adoption are therefore required. It is feasible to amalgamate clinic reports (which will usually be the more reliable though less complete source) with survey information.

Life-table records of contraceptive use-effectiveness may, as discussed earlier, be employed directly for life-table calculations of births averted.

Where a summary mean duration of use is required the formula of Mauldin *et al.* (1967), modified by Potter (1969a), is usually applied. U, the mean span of use is given by:

$$U = \frac{1-X}{P} \tag{3}$$

where
X is the proportion losing the device etc. "immediately" and
P is the (presumed constant) monthly risk of discontinuation.

P is found by locating the value which fits the rate of discontinuation found between specified times after adoption. For example, the rate between the 6th and 24th months is given by:

$$\frac{U_{(24)}}{U_{(6)}} = e^{-18P} \tag{4}$$

While X is derived after the P value is determined from:

$$(1-X)\, e^{-24P} = U_{(24)} \,(^4) \tag{5}$$

In these formulae, $U_{(n)}$ is the proportion retaining the device n months after acceptance.

(4) See footnote 3.

Other, cruder methods have been employed to assess continuation where life-table data are not available. I feel, however, that birth-averted calculations cannot be recommended in the absence of reasonably adequate follow-up information regarding contraceptive use.

The formula

$$U = \frac{1-X}{P}$$

was derived for IUDs, but seems equally applicable to other methods of contraception, though not to sterilization nor, of course, abortion. Until recently, the formula has received surprisingly little critical attention.

One study (Kelly, 1971a) demonstrated that the exponential formula was superior to a reciprocal formula for a (rather small) sample of Puerto Rican acceptors, but found additionally that P was significantly age-dependent. Kelly gave theoretical justification for the exponential formula as follows:

> "Contraceptive patients and their husbands are subject to a number of probabilistic personal events such as illness, death, divorce, migration, conscription, pregnancy, desertion, etc., all of which can take the patient out of contact with the program or remove the need for contraception. If these events can be assumed to be random, then some (constant) value can be assigned to the probability that a patient will continue or discontinue contraception during a given period."

There are, however, very strong theoretical grounds for not assuming that many of these, and other relevant events, are random over time.

The risks of illness and death are continually increasing due

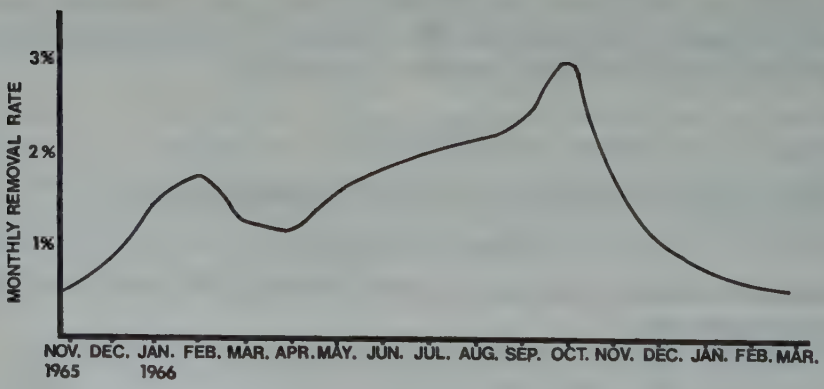


Fig. 6.2. *Removal rate (per unremoved IUD) by date of removal.*

to age, and the risk of pregnancy during use constantly decreases. The likelihood that a user will discontinue in order to have another child undoubtedly rises to a maximum somewhere between two and three-and-a-half years after adoption, so that "spacers" and "terminators" may be expected not only to have different risks of discontinuation, but to form continually changing proportions of all users. Expulsions of IUDs decrease with duration of use; medical complications have well-defined relationships to duration of use. Removals and discontinuations follow the spread of rumours and fashions in a population (see Figure 6.2). For detailed discussion of extrapolated continuation, see Chapter 3.

Estimates of contraceptive prevalence, both within programs and outside them, remain seriously inaccurate in many settings. With definitions of acceptor operationally different in different programs and follow-up requirements neglected or treated casually in many, life-table continuation rates cannot always be relied on, while the difficulties of determining what is happening with non-program contraception, abortion and sterilization in order to assess substitution and trend effects remain essentially insuperable.

MORTALITY, WIDOWHOOD AND DIVORCE

It is a moot point whether separate allowance for mortality, widowhood and divorce should be made for births averted by contraceptive use. The decision must rest upon whether these events are included in the bank of discontinuers in the compilation of use-effectiveness life tables.

On the other hand, for sterilizations (see below) consideration of these events is mandatory. While mortality risks may be taken unmodified from life-tables or age-specific mortality schedules, widowhood, divorce and (if attainable) separation rates must be modified by re-marriage probabilities and intervals to incur a penalty less than the whole of the remaining life-span of the sterilisand.

The practice of combining mortality risks of husband and wife on the basis of their mean age-difference (taken in India as 5 years), as was done by Venkatacharya (1971a) and Ram (1971) as follows:

$$P_x^{mf}(T) = P_x^f(T) \cdot P_{x+5}^f(T) \qquad (6)$$

introduces additional inaccuracy by granting a mean age for all husbands of wives in each age-group. There the mean mortality risk of all the ages is not equivalent to the mortality risk at the mean age.

Mortality and widowhood are usually estimated from life tables, either specific to the population studied or fitted from model tables.

These provide separate annual probabilities of survival from age x to age x+n for males and females.

Divorce statistics are usually inadequate and estimates of the probability of a union being disrupted by divorce or separation during contraceptive use are made from cross-sectional surveys.

DURATION OF PREGNANCY, TIMING OF ABORTIONS AND DURATION OF POST-ABORTAL AMENORRHEA

These are usually arbitrarily set at nine months for a live birth, eight months for a still birth, three months for an abortion and one, one-and-a-half or two months post-abortal amenorrhea.

ACCEPTOR CHARACTERISTICS

All studies place acceptors in age-groups applicable at the time of acceptance. The intervals selected should correspond with the age-intervals of the fertility data to be used ([5])

While some workers have elected to neglect the change in fertility potential due to aging of the acceptor cohort, this cannot be recommended.

Potter (1969a) employs age-specific fertility expectations applicable to the mean age attained by acceptors (in each acceptor age-group) during use, i.e., if mean duration of use is two years, the mean age of acceptors at ages 25-29 will be taken as 28.5 years. Wolfers (1969), by employing single-year age data, ages his population annually.

The division of acceptors into groups according to age, previous contraceptive use, ethnic and social divisions, etc. is justified only insofar as other parameters employed in the calculations are distinguished for these groups. In particular this relates to continuation rates and fertility expectations. Separate runs of calculations will be required for each distinguishing group.

Other characteristics of acceptors, notably open interval to acceptance, are directly integrated into the calculations as described elsewhere.

Characteristics of contraceptive users are obtained (preferably) from service statistics where, at the time of acceptance, necessary items such as age, previous contraceptive use, social characteristics and open interval are recorded. While age will usually be computed

([5]) If the formula on page 19 is employed, acceptors should therefore be grouped in ages 18-22, 23-27, 28-32, etc.

in 5-year groupings, care needs to be taken with mean age within groups at the extreme ends of the child-bearing period. In the Singapore post-partum program, the mean age of acceptors 15-19 years old was 18.9 years—1.4 years higher than the mid-point of the interval; at 20-24 it was 23.0 years, at 40-44, 41.7 and at 45-49, 46.5 years (Wolfers, 1970a: 169).

Alternatively, this information, but with considerable loss of accuracy, may be collected by retrospective sample surveys.

INDUCED ABORTION

Somewhat different modes of calculation are required to compute the birth-averting powers of induced abortions and sterilizations.

Induced abortions, not being continuously employed strategies, do not lend themselves to the calculations of duration of use. Instead it is necessary to seek a quantity, "duration per abortion", to correspond to "duration per birth". This, like the common birth interval, will be composed of three parts: a period of ovulatory exposure, a period of pregnancy and a period of post-partum anovularity. The latter two quantities are subject to little variation and have mean values close to 3 months and $1\frac{1}{2}$ months respectively. The former may be derived from fecundability estimates, but these will not usually be available and so will commonly be "borrowed".

"Duration per abortion" now becomes the period corresponding to mean duration of contraceptive use, and its birth-averting effect is obtained from the formula:

$$\frac{\text{Duration per abortion}}{\text{Duration per birth}} = \text{births averted per abortion} \qquad (7)$$

(Potter, 1972b).

While this approach yields a satisfactory solution to the problem of attributing a birth-preventing value to abortions where they are the only form of birth control employed, this is not the common case. Usually abortion is used to supplement a more conventional method of contraception.

Under these circumstances, the duration per birth is that to be expected with the use of the contraceptive and the duration per abortion is correspondingly increased because of the elongation of the period of ovulatory exposure.

The appropriate duration per birth may be obtained by summing the duration per birth used in the calculation of births averted by contraceptive use for the age-group concerned and the mean prolongation of stay in the fecundable period calculated on the basis of their contraceptive use (see section IV, 2).

In respect of abortion, the necessity of taking post-period strategy into account is most clearly demonstrated. The impact of abortion in averting births has been extensively discussed by Potter (1972b) and Keyfitz (1971). Potter has also made a more theoretical mathematical exploration (1972c).

STERILIZATION

While there are serious substitution problems in computing births averted by sterilizations—the probability that sterilisands would have had recourse to other methods of birth control had sterilization not been available is very high—it is probably a healthy sign that they have been almost entirely ignored by writers on this subject. By the time we have taken into account that, but for sterilization, half our acceptors would have worn IUDs and but for IUDs half these acceptors would have been sterilized, our overscrupulousness is thrusting half the program into limbo!

Births averted by sterilization comprise all the remaining potential births to the couple involved. The important dimensions of the problem of quantification are:

a) the rates to be used—all women or past fertility of acceptors only;

b) ages at acceptance—and in this case, the use of mean age of acceptors is not satisfactory (see above);

c) mortality, widowhood and separation (see section III.7 above);

d) susceptibility status (physiological state) at time of sterilization.

The majority of studies of births averted by sterilization have been made using Indian experience. Gopalaswami (1962) based his calculations for India on Japanese fertility data, Agarwala (1966) employed ASMFRs, while Haynes *et al.* (1969) matched sterilisands in Kerala with controls by age, income, education, religion and number of living children and used the fertility rates of controls as their basis—finding a 10 per cent increase over ASMFRs estimated for the same period. Venkatacharya's calculations (1971a, 1972) have as their basis the ASMFRs of the 1954-55 Mysore study, for which two alternative sets of fecundability schedules are derived. Were they but available, the use of age-controlled parity progression ratios would provide an excellent theoretical basis for estimating births averted by sterilization.

Venkatacharya (1971a) claiming that almost all female steriliza-

tions in India take place immediately after a delivery, makes specific allowance for a lengthy period of non-susceptibility in risk of conception at the outset of sterilizations, and a moderated allowance for vasectomies.

Future fertility estimations for sterilizations must take into account the low proportion of sterile acceptors but, this aside, it is doubtful if higher fecundity than average should also be assumed.

Even the best estimate of the effects of sterilization encounter the difficulty that effects are continuing for as long as two decades after the act, so that the validity of assumptions about expected fertility rates is always questionable.

SUBSTITUTION

In practice substitution must be handled by making the assumption that a certain (always to some extent arbitrary) proportion of acceptors is using the program method as a substitute for a non-program method, and assigning to them a lower potential fertility level (or zero potential fertility) than to the remainder.

The probability that acceptors in the program *would have* adopted non-program birth control strategies were it not for the program has also been addressed.

Potter (1969a) calculated alternative D-values (durations per birth) on differing assumptions regarding the probability that acceptors would have been sterilized if IUDs were not available, and Wolfers (1969) attempted to calculate the monthly age-specific proportion of married women becoming successful contraceptors outside the program and amalgamated this with sterility probabilities.

While at the heart of the question of how much realism resides in birth-averted calculations, the quantification of substitution effects remains elusive and becomes more so as societies develop, motivation for birth control grows and channels of supply of services multiply. In universally contracepting populations every episode of use is a substitute for some alternative that was or would have been used instead.

Of special importance in the case of abortions, but also relevant to the assessment of the effects of contraceptive use is the question of birth-control strategy to be employed after discontinuation of a program method.

Whether the concept of extended use-effectiveness of Tietze and Lewit (1968) should be employed to credit a program with the catalytic effect it has in keeping acceptors contracepting after leaving the program, or whether recognition of the fact that discontinuers have low fertility expectations should modify downward the estimates

we make of births averted, will depend on our purpose in making the calculations—essentially a political, rather than a scientific, decision.

Substitution estimates will be based on data collected at the time of acceptance and survey information regarding duration and efficacy of contraceptive practice outside the program. The accuracy of the latter will rarely be great.

CHOICE OF BASELINE AND TREND

Calculations of the demographic effects of programs, whether by births averted or other computations, have been attacked, for example by Seltzer (1970), for failing to incorporate realistic assumptions as to what national fertility would have been without the program. Apart, therefore, from the substitution problems which exist in steady-state societies, there are others, yet more speculative, which derive from trends already manifest at the time of introduction of the program. It is desirable, where sufficient data exist to do so, to incorporate assumptions based on these trends at least conditionally, or as alternatives, into computations of births averted. This will not entirely disarm criticism, for any particular sequence of falling fertility rates can be extrapolated in many different ways, but some extrapolation is desirable.

The simplest extrapolation is to assume an arithmetic rate of fall—for example, if ASMFR for age 20-25 has fallen by 50 per thousand from 350 to 300 per thousand in the past five years, to assume that it would have continued to fall at the rate of 10 births per thousand per year. More realistic is to assume a geometric or logarithmic rate of fall of 3 per cent per year. Other suggestions include allowance for an accelerating rate of fall. One method used by Wolfers (1970b) (although not in births-averted calculations) is first to compute an approximation of an ASMFR schedule for the population which corresponds to zero growth rate and then using this as the zero fertility level, render ASMFRs into "excess" fertility rates and find trends geometrically from there. Thus in the example given, if the zero growth ASMFR for age 20-25 is 150, the fall in "excess" fertility rate will have been from (350-150) to (300-150) or 25 per cent over five years, equivalent to a fall of 5.6 per cent per year.

Trends of this kind have not, to my knowledge, so far been incorporated into birth-averted calculations, except insofar as is possible in considering second derivative substitution effects, but it is desirable that they should be.

IV. ILLUSTRATED EXAMPLES

KOREA: LEE AND ISBISTER'S METHOD

Although earlier attempts to quantify the effect of contraceptive use do exist, the first full-scale computation of births averted was that of Lee and Isbister (1966), under the guidance of Ansley Coale. This pioneering calculation had the comprehensive aims of estimating:

a) the effects of a program on the fertility of individual future years;

b) the total effect that IUDs will have during all the years that they remain in use;

c) the scale of program required to achieve specified objectives.

The program used in the illustration was that of Korea and the method employed was essentially that of component projection, with future age-specific fertility rates calculated by using the formula:

$$f_{i,t} = \frac{F_{i,t} \cdot f_{i,0} - Q_{i,t}\, g_i}{F_{i,t}} \qquad (8)$$

where,

$f_{i,t}$ is the fertility of women aged i in year t ($t = 0$ for base year);
$F_{i,t}$ is the total number of women aged i in year t;
$Q_{i,t}$ is the number of women aged i in year t practising totally efficient contraception in year $t-1$ (on the assumption that contraceptive use averts births approximately one year later);
g_i is the potential fertility of women aged i on the assumption of non-use of (program) contraception. As indicated above (p. 177), g_i was taken in this study as $f_{i,0} + 20\ \%$.

Perhaps the most serious flaw in this method is that it takes no account of the initial "susceptibility status" (physiological status) of the women comprising the $Q_{i,t}$ group, leading to under- or over-estimation of future fertility rates depending on whether contraception was being adopted during post-partum amenorrhea or during the fecundable state.

Life-table methods of computing contraceptive use-effectiveness were not yet available when Lee and Isbiter formulated their proposals so that their method of calculating the size of the Q group leaves much to be desired. Anyone wishing to employ the method nowadays

would employ life-table methods to calculate the numbers using contraceptives in specific years from annual data.

The method employed for determining births prevented in all years per IUD inserted is a simple adaptation of the fertility rate calculation, the survivors wearing IUDs being multiplied by appropriate g_i potential fertility rates until all have discontinued use.

Lee and Isbister computed that each IUD inserted prevented between 1.5 and 1.9 births, depending on the age distribution of acceptors.

Two very similar methods of calculating births averted were published in 1969, one by Potter, using Taiwan data, and one by Wolfers employing data from the post-partum IUD program in Singapore. Both endeavored to take into account a wider range of factors than were recognized in Lee and Isbister's work.

TAIWAN: POTTER'S METHOD

The rationale of Potter's method is expressed as follows:

> "For an estimate of births averted per segment of IUD, two kinds of statistics are needed. The first are estimates of the prolongations of stay in the fecundable state resulting from retention of the device. Second, to convert these prolongations into births averted it is necessary to divide them by a constant representing the average marriage duration per birth that might have been required by the couples had they not adopted IUD."

Two formulae are used:

$$I = F(R - A - PW) \tag{9}$$

where

I = mean prolongation of stay in the fecundable period;
F = proportion of acceptors non-sterile at time of insertion;
R = mean time device is "retained" among couples fertile at time of insertion;
A = allowance for amenorrhea;
P = proportion becoming pregnant while believed to be retaining the device;
W = penalty for accidental pregnancy;

$$B = I/D \tag{10}$$

where

B = births averted per first segment of IUD;
D = average duration per birth.

Amplifying,

R, the mean time of "retention" is defined as the mean interval

from insertion to death, widowhood, divorce, onset of secondary sterility or loss of contraceptive effect through pregnancy, expulsion or removal, whichever occurs sooner. R is estimated by the technique of multiple decrement life-table analysis with each eventuality treated as a competing risk. Note, however, that where removal is ascertained by follow-up survey, and widowhood and divorce from population registration statistics, duplication may occur if care is not taken.

A is the mean overlap between use of IUD and post-partum amenorrhea, obtained as the weighted mean of the results of subtracting delivery-insertion intervals, grouped in three-month classes, from mean duration of amenorrhea which is taken (Mohapatra, 1966) as 8 months.

P is the proportion of acceptors fertile at insertion whose useful span of protection is terminated by pregnancy.

W is the mean fecundable period that would have been required for a pregnancy in the absence of an IUD. (As Potter points out, were IUDs substituting for more effective contraception, the IUD would in fact be shortening average stays in the fecundable period, the value of W would be high and PW would exceed R-A, exposing the fact that I, the prolongation of stay in the fecundable period, has a negative value and births are being encouraged rather than averted by the contraceptive program.)

Potter obtains births averted by years after insertion by subdividing I, the interruption of child-bearing, into sub-totals relevant to each year after insertion and dividing each of these by D. He does not recommend varying D as the remaining users age, holding that D changes (except towards the end of reproductive life) too slowly to justify such variation.

The mode of calculating proportions sterile at insertion and those becoming sterile during use is described in section III.5.

In calculating the mean span of useful retention of the IUD, Potter amalgamates the competing risks of pregnancy, expulsion and removal (PER) and secondary sterility and mortality (of either spouse). The method of handling PER risks is that described earlier, assuming a proportion losing the device immediately and a constant proportion of remaining users losing it in each subsequent month.

Mortality is treated by selecting approximations for the average death rates for each age class. (For ages 20-29, the death rates (r) in those classes are used; for ages 30-34 and 35-39, death rates for 30-39 and 35-44 are employed.) Based on these probabilities of survival for one year are approximated by $1-r$ and designated P_f and P_m for wives and husbands respectively. The monthly risk of death *or* widowhood is then given by M,

Table 6.2. *Computation of cumulative births*

(a) Ordinal month	(b) Surviving fecund users at 1st of month	(c) Month after delivery	(d) Probability integral	(e) Mean post-partum infecundity	(f) Effective population	(g) Continuation rate to end of month	(h) Continuation rate × modifying factor	(i) (h) × sterility factor
1	96668	1.5	0.0316	0.6	3055	0.918495	0.911396	0.908313
2	87805	2.5	0.0580	1.3	5093	0.970764	0.963261	0.960002
3	84293	3.5	0.0993	2.0	8370	0.972004	0.954491	0.961229
4	81025	4.5	0.1587	2.8	12859	0.958324	0.950917	0.94770
5	76787	5.5	0.2376	3.5	18245	0.974727	0.967103	0.963921
6	74017	6.5	0.3341	4.2	24729	0.96659	0.959126	0.955881
7	70751	7.5	0.4432	4.9	31357	0.974867	0.067332	0.964060
8	68208	8.5	0.3568	5.5	37978	0.974213	0.966683	0.963413
9	63712	9.5	0.6659	6.1	43758	0.977093	0.969541	0.966261
10	63495	10.5	0.7624	6.6	48409	0.978182	0.970622	0.967388
11	61421	11.5	0.8413	7.0	51673	0.987720	0.980086	0.976770
12	59994	12.5	0.9007	7.3	54037	0.980451	0.972873	0.969382
13	58169	13.5	0.9420	7.6	54795	0.988470	0.980830	0.977512
14	56881	14.5	0.9684	7.8	55064	0.989173	0.981528	0.978207
15	55622	15.5	0.9839	7.9	54726	0.990603	0.982847	0.979522
16	54483	16.5	0.9925	7.9	54074	0.989696	0.982047	0.978724
17	53324	17.5	0.9966	8.0	53143	0.992706	0.985033	0.981701
18	52348	18.5	0.9987	8.0	52280	0.995845	0.988148	0.984805
19	51553	19.5	1.000	8.0	—	0.990164	0.982511	0.979187
20	50486		1.000	8.0	—	0.993	0.987310	0.983970
21	49671		1.000	8.0	—	0.993	0.987310	0.983970
22	48875		1.000	8.0	—	0.993	0.987310	0.983970
23	48092		1.000	8.0	—	0.993	0.987310	0.983970
24	47321		1.000	8.0	—	0.993	0.987310	0.983970
25	46562		1.000	8.0	—	0.993	0.987310	0.991634
26	46178		1.000	8.0	—	0.993	0.987310	0.991634
27	45786		1.000	8.0	—	0.993	0.987310	0.991634
28	45403		1.000	8.0	—	0.993	0.987310	0.991634
29	45033		1.000	8.0	—	0.993	0.987310	0.991634
30	44646		1.000	8.0	—	0.993	0.987310	0.991634

The above table is an example for Singapore for the Chinese ethnic group aged 25-29 with no previous contraceptive history, and using the following factors:

averted per 100,000 women accepting contraception

(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)
Women-months of use in month	Corrected mean birth-interval	Births averted (100 % effect)	Probability of pregnancy	Probability of pregnancy × live-birth prop.	Births	Corrected births averted	Cumulative births averted	Number sterile or discontinued in month
2915	17.8	164	0.001135	0.000793	79	63	85	8063
4991	18.5	270	0.002270	0.001590	159	111	196	3518
8207	19.2	427	0.001702	0.001191	119	308	304	3238
12522	20.0	626	0.002459	0.001721	172	454	958	4338
17915	20.7	865	0.001891	0.001324	132	733	1691	2770
24183	21.4	1130	0.001891	0.001324	132	993	2632	3266
30793	22.1	1393	0.002269	0.001583	150	1234	5983	3502
37283	22.7	1642	0.003215	0.002251	225	1417	5340	2493
43019	22.3	1646	0.002648	0.001854	185	1661	7001	2217
47618	23.8	2001	0.001136	0.000795	80	1921	8922	2074
51073	24.2	2110	0.001707	0.001195	119	1991	10913	1427
53215	24.5	2172	0.001905	0.001334	133	2039	12952	1325
54179	24.8	2185	0.000774	0.000542	54	2391	15083	1898
54464	25.0	2179	0.001840	0.001288	129	2630	17183	1238
54166	25.1	2158	0.000710	0.000497	50	2198	19341	1109
53499	25.1	2131	0.000312	0.000218	22	2109	21350	1159
52656	25.2	2090	0.001924	0.001347	135	1855	23903	976
51950	25.1	2062	0.000000	0.000000	0	2062	25267	733
51017	25.2	2024	0.000000	0.000000	0	2024	27391	1073
50076	25.2	1987	0.0006	0.00042	42	1945	20836	909
49273	25.2	1955	0.0006	0.00042	42	1913	31269	790
48485	25.2	1924	0.0006	0.00042	42	1862	38131	783
47701	25.2	1893	0.0006	0.00042	42	1851	34088	773
46942	25.2	1863	0.0006	0.00042	42	1881	36963	159
46367	25.2	1840	0.0005	0.00035	35	1885	38608	290
45979	25.2	1825	0.0005	0.00035	35	1780	40298	380
45594	25.2	1809	0.0005	0.00035	35	1774	48172	288
45213	25.2	1794	0.0005	0.00035	35	1732	43081	280
44835	25.2	1779	0.0005	0.00035	35	1744	45075	377
44439	25.2	1764	0.0005	0.00035	35	1783	47464	278

mean prospective birth interval (A) = 25.2 months; continuation modifying factor (B) = 0.992271; withdrawal factor (C) = 0.293653.

Table 6.2. *Computation of cumulative births averted*

(a) Ordinal month	(b) Surviving fecund users at 1st of month	(c) Month after delivery	(d) Probability integral	(e) Mean post-partum infecundity	(f) Effective population	(g) Continuation rate to end of month	(h) Continuation rate × modifying factor	(i) (h) × sterility factor
31	44872		1.000	8.0	—	0.993	0.987310	0.991634
32	43902		1.000	8.0	—	0.993	0.987310	0.991634
33	43535		1.000	8.0	—	0.993	0.987310	0.991634
34	43171		1.000	8.0	—	0.993	0.987310	0.991634
35	42810		1.000	8.0	—	0.993	0.987310	0.991634
36	42462		1.000	8.0	—	0.993	0.987310	0.991634
37	42097		1.000	8.0	—	0.993	0.987310	0.991634
38	41745		1.000	8.0	—	0.993	0.987310	0.991634
39	41396		1.000	8.0	—	0.993	0.987310	0.991634
40	41050		1.000	8.0	—	0.993	0.987310	0.991634
41	40707		1.000	8.0	—	0.993	0.987310	0.991634
42	40366		1.000	8.0	—	0.993	0.987310	0.991634
43	40028		1.000	8.0	—	0.993	0.987310	0.991634
44	39692		1.000	8.0	—	0.993	0.987310	0.991634
45	39360		1.000	8.0	—	0.993	0.987310	0.991634
46	39031		1.000	8.0	—	0.993	0.987310	0.991634
47	38704		1.000	8.0	—	0.993	0.987310	0.991634
48	38380		1.000	8.0	—	0.993	0.987310	0.991634
49	38052		1.000	8.0	—	0.993	0.987310	0.991634
50	37741		1.000	8.0	—	0.993	0.987310	0.991634
51	37425		1.000	8.0	—	0.993	0.987310	0.991634
52	37112		1.000	8.0	—	0.993	0.987310	0.991634
53	36802		1.000	8.0	—	0.993	0.987310	0.991634
54	36494		1.000	8.0	—	0.993	0.987310	0.991634
55	36189		1.000	8.9	—	9.993	0.987310	0.991634
56	35880		1.000	8.0	—	0.993	0.987310	0.991634
57	35680		1.000	8.0	—	0.993	0.987310	0.991634
58	35838		1.000	8.0	—	0.993	0.987310	0.991634
59	34993		1.000	8.0	—	0.993	0.987310	0.991634
60	34700		1.000	8.0	—	0.993	0.987310	0.991634

per 100,000 women accepting contraception (continued)

(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)
Women-months of use in month	*Corrected mean birth-interval*	*Births averted (100 % effect)*	*Probability of pregnancy*	*Probability of pregnancy × live-birth prop.*	*Births*	*Corrected births averted*	*Cumulative births averted*	*Number sterile or discontinued in month*
44087	25.2	1749	0.0004	0.00028	28	1721	49125	270
43718	25.2	1735	0.0004	0.00028	28	1707	39858	387
43353	25.2	1720	0.0004	0.00028	28	1092	48524	304
42980	25.2	1706	0.0004	0.00028	28	1678	54808	261
42631	25.2	1698	0.0004	0.00028	28	1865	55866	252
42274	25.2	1678	0.0004	0.00028	28	1683	57516	501
41921	25.2	1654	0.0002	0.00014	14	1658	59165	382
41370	25.2	1650	0.0002	0.00014	14	1630	60808	340
41923	25.2	1636	0.0002	0.00014	14	1682	62484	348
40878	25.2	1623	0.0002	0.00014	14	1608	64932	340
40537	25.2	1609	0.0002	0.00014	14	1595	63827	361
40197	25.2	1595	0.0002	0.00014	14	1581	67398	328
39861	25.2	1582	0.0002	0.00014	14	1568	68776	335
39327	25.2	1569	0.0002	0.00014	14	1555	70321	388
39195	25.2	1565	0.0002	0.00014	14	1541	71272	329
38868	25.2	1542	0.0002	0.00014	14	1528	70406	327
38542	25.2	1529	0.0002	0.00014	14	1515	74913	384
38310	25.2	1517	0.0002	0.00014	14	1505	70418	261
37000	25.2	1504	0.0002	0.00014	14	1490	71290	318
37683	25.2	1491	0.0002	0.00014	14	1477	79384	210
37268	25.2	1479	0.0002	0.00014	14	1468	80650	414
36957	25.2	1467	0.0002	0.00014	14	1453	82898	210
36648	25.2	1484	0.0002	0.00014	14	1440	83760	360
36341	25.2	1442	0.0002	0.00014	14	1428	86371	208
36038	25.2	1430	0.0002	0.00014	14	1416	86567	293
35736	25.2	1418	0.0002	0.00014	14	1404	87901	390
35457	25.2	1408	0.0002	0.00014	14	1398	88383	290
35140	25.2	1394	0.0002	0.00014	14	1380	90763	280
34867	25.2	1383	0.0002	0.00014	14	1369	92182	302
34555	25.2	1371	0.0002	0.00014	14	1387	93689	289

where

$$P_f P_m = e^{-12M} \tag{11}$$

The use of three exponential expressions to denote the competing risks of PER (p), mortality (m) and sterility (s), together with the convention that a proportion, x, lose the device immediately, allows the estimation of the mean retention span R to be expressed by the simple expression

$$R = \frac{1-x}{m+s+p} \tag{12}$$

The proportions of EUROD attributable to each of the competing causes, m, s and p are then

$$\frac{(1-x)\,m}{m+s+p}, \quad \frac{(1-x)\,s}{m+s+p} \quad \text{and} \quad \frac{x+(1-x)\,p}{m+s+p}$$

This last becomes useful as a basis for calculation of P, the proportion accidentally pregnant. P (age-specific) is obtained by multiplying this proportion of loss of useful IUD effect due to pregnancy, expulsion or removal by the ratio of accidental pregnancy to all events in the use-effectiveness life table. However, as the ratio of pregnancy to all events changes as a function of duration of use, Potter recommends a rather elaborate process to obtain an accurate estimate. The ratio, pregnancies to all events, is divided into R_1 and R_2, R_1 being the ratio for the first two years, R_2 for the subsequent life of the IUDs, with the overall ratio obtained by a weighted average of R_1 and R_2. Weighting is attained by total durations of use in those two periods. R_1 is obtained directly from the life table; R_2 is calculated on the assumption that no further expulsions occur after two years and that pregnancies form a duration-independent proportion of pregnancies and removals.

This extrapolation is rather artificial but is dictated by the absence of long-term life-table data. Its validity is heavily dependent on the reality of the formula for exponential decay in useful retention.

Calculation of penalty per pregnancy was described in Section III.

The basis for the calculation of durations per birth was, as discussed above, the pre-insertion birth rates of acceptors obtained by survey. Birth rates for the three years preceding the first interview were used. Meticulous care was required to ensure that appropriate time relationships were employed. The point of general relevance is that it is essential to employ the fertility rates appropriate to women of the ages that acceptors will have reached at the mid-point in time of their useful retention of IUDs. Durations per birth are derived from:

$$D_i = \frac{12{,}000}{b_i} \tag{13}$$

where
D_i is the duration at age i
and
b_i is the birth rate at that age.

Potter calculates a series of four alternative D-values employing different assumptions regarding the probability that IUD acceptors would have become sterilized if IUDs were not available.

Equation (10) above is then applied to yield births averted per IUD inserted. The method, when applied to Taiwan data, yielded estimates of this value ranging from 0.94 to 0.43, depending on which of the four alternative D values was used.

SINGAPORE: WOLFER'S METHOD

In contrast to Potter's method, which is characterized by the use of summary values for duration of use, post-partum amenorrhea and pregnancy rate, Wolfers employs a life-table approach to yield, in each ordinal month after acceptance, the proportions of original acceptors who are still usefully retaining the device and with what effect.

Data and calculations are assembled in a series of tabulations of the form shown in Table 6.2. In the example worked, separate tables are constructed for each population sub-group by age, ethnic group and previous contraceptive use. The three factors, (A), (B) and (C) in the note of the table are the distinguishing variables for sub-groups, and are obtained as follows:

(A) The mean prospective birth interval is obtained from direct questioning of the population, acceptors and non-acceptors, and is, as described above, a mean interval for births. It corresponds to Potter's D-values.

(B) The continuation modifying factor is a factor applied to life-table continuation rates computed by the method of Tietze and Potter in an attempt to make these specific for sub-groups. The sub-group proportions found to be continuing use in a follow-up survey some sixteen and a half months (on average) after acceptance were applied as correction factors to the general life-table rates. Thus while the proportion of the total population continuing use at 16.44 months was 0.66, the proportion of Malays aged 35-39 continuing use was 0.88 or 1.33 times as high as the average. The 16.44th root of this ratio - 1.01765 was then applied to each monthly continuation rate (up to the 24th month) to achieve a rate specific for that group.

(C) The withdrawal factor. Two alternative values are computed for this variable:

(i) represents only the monthly probability of the development of sterility (see section III.5), while

(ii) combines the probability of becoming sterile with the probability of ceasing to reproduce for other reasons and is derived from estimates made of the proportion of the population at each age, still engaged in reproduction. It represents an attempt to include second derivative substitution effects. These estimates are made by calculating an expected number of births for each age-group based on numbers married, proportion of first births, mean *retrospective* birth interval and estimated marriage-first birth interval. The formula employed is:

$$1.25B_1 + IB_+ = \text{Number of women married 9 months and longer,} \qquad (14)$$

where

B_1 is the number of first births

B_+ is the number of later births

I is the mean retrospective birth interval.

With knowledge of the ratio B_1/B_+, this equation can be solved to give the expected total births on the basis of the known married population. The actual/expected ratio gives the proportions of women in each age-group still reproducing and again, the 60th roots of five-year rates of change yield a monthly proportion not withdrawing from reproduction.

The calculation of births averted proceeds month by month as follows:

A notional cohort of 100,000 acceptors is first reduced to the number fertile at acceptance on the assumption that all were fertile at the *conception* of the last child (cf. Potter's assumption). Monthly probabilities of sterility (or withdrawal) are applied for nine months plus the mean length of the delivery acceptance interval. Thereafter, each subsequent month of use is treated similarly. For the nth post-acceptance month a computation is made of the proportion of the still fertile acceptors who are *not* experiencing post-partum amenorrhea in that month. The mean duration of post-partum amenorrhea for each sub-group was estimated by the formula

$$\overline{F} = 0.9\left(2\,(V/8)^{\frac{1}{2}} + 1\right) \qquad (15)$$

where

V is the total variance of birth-intervals in the sub-group

$\overline{F}$ is the mean post-partum amenorrhea.

Making the simplifying approximation that post-partum amenorrhea was normally distributed, and that the mean amenorrhea was twice the standard deviation plus one month, the proportions still amenorrheic in each month could be read from tables of probability integrals (areas under a curve to the left of each ordinate value). The use of an observed distribution, if available, would, of course, have been preferable.

By multiplying fertile acceptors by this integral, the number of women in the cohort who were neither sterile nor amenorrheic was obtained—the effective population. This number was then serially multiplied by the continuation rate to the end of the month (from the use-effectiveness life table), continuation modifying factor (B) and sterility or withdrawal factor (C) to yield continuing fertile users from this group at the end of the month. Adding half the month's decrement to this number gave woman-months of effective contraceptive use during that month.

Wolfers then argued that an earlier return of menstruation pretold shorter birth-intervals, assuming that the duration of post-partum amenorrhea remains fairly constant from birth to birth for any one woman (but see Barrett (1971) who takes a different view) and corrected the birth interval for the sub-groups by subtracting from it the mean post-partum infecundity of those women who had already resumed menstruation before the middle of the nth month. These values are also obtainable from probability integrals. This subtraction yields corrected intervals. Births averted by contraceptive use in the nth month are then obtained by dividing woman-months of use by the corrected mean birth interval and then subtracting the pregnancies occurring (obtained as 100,000 $\times$ pregnancy rate for month/n from life table) multiplied by the proportion of pregnancies terminating in a live birth (taken as 0.7). This method, by which pregnancies effectively carry a penalty equivalent to the full live-birth interval, is less accurate or justified than Potter's method.

The surviving fecund users for the next month $(n+1)$ are then obtained by multiplying the surviving fecund users for month n serially by the continuation rate for month n, the continuation modifying factor (up to 24 months only) and the withdrawal factor.

After a varying period, some of the columns in the table stabilize; i.e., the probability integrals tend towards a value of 1 indicating that all women have by then emerged from post-partum amenorrhea. When this value is reached, the effective population thereafter is equal to the surviving fecund users. The mean infecundity of the population already ovulating tends towards the mean infecundity of the whole

population. When this value is reached, no further corrections of mean birth interval are indicated.

Continuation rates cease to vary after the limit of life-table evidence is reached and continuation modifying factors are dropped at 24 months, but recent evidence (e.g. Avery, 1973) suggests that these procedures could be improved.

Wolfers allowed pregnancy rates to decline up to three years after acceptance and then to level off. An improvement would have been to stipulate that after life-table experience was exhausted pregnancy rates—steady or declining—should be applied only to surviving fecund users instead of the cohort of initial acceptors.

Considerable attention is paid by Wolfers to the changing value of parameters as the population of effective users ages, and tabulations are made of the size of each sub-group for each year studied; single year age data are used in the reclassification. As users move from one age-group table to the next, transfer factors are employed to compensate for the differences in numbers of surviving effective users in the same month, n, in different age-groups. Thus, in the most extreme case, after 4 years' use there are still some 48,000 surviving fecund Malay users aged 40-44. When transferred to the table for women aged 45-49, they enter a line with only 3,458 survivors of 100,000 original entrants. The number transferred is therefore multiplied by 14.

Wolfers found in Singapore that one IUD averted (for the first three years after insertion) 0.57 births for Chinese women, 0.64 for Indian women and 0.79 for Malay women. Ethnic differences reflected higher fecundity, higher continuation rates and lower previous practice of contraception in the Malay groups. The weighted average was 0.59.

Wolfers' method is designed to yield births averted, by years after acceptance, in a program. It does not lend itself readily to computing births ever averted per contraceptive supplied.

COMPUTER SIMULATION

A fourth method of calculating births averted, one which in my view is likely ultimately to supplant all those described above (and the numerous variants to which reference has or has not been made) is the use of computer models to simulate the reproductive histories of cohorts of women from marriage to cessation of reproduction. Any desired family planning strategy can be introduced into such models and the results, in terms of births averted both in relation to items of contraceptive use and to time, can be derived by comparison with

reproductive performance in the absence of the strategy. Ridley and Clague (1972) have demonstrated the technique in an intriguing experiment in which they tested each of the three methods described above—Lee and Isbister, Potter and Wolfers—against computer estimates obtained in this way. Family planning practice in the model was restricted to women over 35 so that the test was applied to women close to the end of reproductive life where truncation effects and assumptions regarding sterility and differentials between fertility of acceptors and the general population begin to assume major proportions in relation to results.

Over a five-year program period, Lee and Isbister's method overestimated births averted by 32 per cent and Wolfers' method by 8 per cent, while Potter's method underestimated by approximately 13 per cent. It is probable that had the test been performed on acceptors between 25 and 34, all three methods would have yielded results closer to those obtained by the computer and to each other's. The use of simulation models for direct computation of births averted in live programs, as distinct from their use as tools for the exploration of theoretical relationships, is handicapped at present by the scarcity of elaborated programs, workers available to employ them, and computer time.

In addition to the use of models made as realistic as possible by the selection of the most suitable parameters and the incorporation of all factors influencing reproduction, including death, divorce, widowhood, sterility, foetal wastage, etc., for which we are indebted to the work of Sheps and Perrin, Brass, Barrett, Potter, Ridley and Venkatacharya, another type of model, based on pure renewal theory, has been employed by Potter to explore the theoretical relationships involved in the calculation of births averted. This approach has been fruitfully employed in analyzing the effects of abortion (Potter, 1972c), the timing of IUD insertions (Potter *et al.*, 1973) and substitution effects (Potter, 1969b).

KELLY'S METHOD

W. J. Kelly (1971b) uses a method which computes the mean annual change in parity by age of acceptors—a numerator analysis approach. Because of the vastly differing probabilities for women of common age but differing parity becoming acceptors, it is difficult to accept that this value has any constant relationship to potential fertility.

OTHER

Apart from Venkatacharya's method, to which reference has been made in connection with the timing of births averted (section II.11), many other methodologies have been used, but these have mainly been either simplifying modifications of one or other of the methods described above or approaches seeking to solve somewhat different problems from that implied by the definition provided.

V. SYNTHESIS

It is feasible to combine some of the better features from these several approaches into a common system, more accurate and possibly simpler to employ, than any.

Every system needs to take into account:

a) "duration per birth" for each population sub-group;
b) use-effectiveness life-table continuation rates;
c) pregnancy rates and penalty per pregnancy;
d) proportion of acceptors sterile at acceptance;
e) development of sterility during use;
f) post-partum anovularity overlap;
g) substitution estimates.

Additional refinements, such as correcting durations per birth for shorter characteristic periods of post-partum amenorrhea, the consideration of mortality and widowhood, etc. separately from discontinuation rates, can safely be omitted without introducing serious inaccuracy.

Duration per birth cannot, in general field use, be calculated either from past fertility of acceptors or from birth intervals without costly and extensive surveys. A simpler derivation from available statistical estimates is essential for family planning programs without major research resources. Age-specific marital fertility rates, whether directly computed or estimated from age-specific fertility rates, will inevitably form their basis.

We would suggest the use of these rates modified as in equation (1) (page 178). This formula assimilates the development of sterility during use, and thus obviates the need for separate consideration of this possibility.

Life-table continuation rates may be used either in life-table form or converted into mean duration of use, depending on which general method of computation is selected.

Monthly pregnancy rates for periods covered by observation will be used, but their extrapolation beyond this period is a matter of judgement. Potter's suggestion for dealing with this appears to be somewhat cumbersome and it might be preferable to make an arbitrary estimate of the monthly risk of pregnancy related to the remainder of the population usefully employing the contraceptive during that month, and consistent with the life-table values.

Penalty per pregnancy should be computed, as described by Potter (see section III.3). In adapting the "penalty per pregnancy" concept to life-table use, an additional step between columns (n) and (o) of Table 6.2, in which the product of pregnancy rate, live birth ratio and population is further multiplied by penalty per pregnancy and divided by duration per birth, (A), or column (k), if it is employed, is required.

Whether arithmetic, logarithmic or exponential decay curves are employed in the calculation of monthly sterility risk makes little substantial difference except at ages over 40. Considerable difference does, however, result from the decision to include or exclude the risk of development of sterility incurred during the nine months of the last preceding pregnancy. Hard evidence to justify a recommendation is lacking, but theoretical considerations lead to the view that the probability of the development of permanent sterility is at least as high during an interval of time which embraces a pregnancy and live birth as during an equal interval which does not. Vincent (1950) and Henry (1953a) both consider that pregnancy imposes a sterility risk *additional* to the normal risk associated with aging. Henry (1953b) was able to confirm this for Norwegian women, but not for Japanese. I suggest, therefore, that proportions fertile at acceptance be calculated by applying monthly sterility risks to the period from the last conception to acceptance, i.e., open interval plus nine months.

Post-partum anovularity is best processed as described in Section III.2. If data are lacking, the process should nonetheless be performed on an artificial distribution of anovularity duration ([6]) with a mean consistent with whatever information about the population is available.

Adopters who are reported, in service statistics, as currently practicing other forms of contraception at the time of adoption may be assigned longer durations per birth than others. The modification should reflect an informed estimate of the likelihood that these couples would have continued to use the other method and for what

([6]) Potter (Potter and Masnick, 1971) follows Barrett (1969) in his most recent papers in assuming a Pascal distribution for post-partum amenorrhea.

mean length of time, as well as the use-effectiveness of the contraceptives involved. Very crudely, this could be achieved by calculating roughly the prolongation of stay in the fecundable state attributable to the non-program contraceptive and subtracting this from prolongations of stay achieved by the program contraceptive.

Thus, if it is estimated that a proportion, p, of current users would have continued to use non-program methods for a mean period of m months with a failure-rate of r pregnancies per 100 woman-years of use, a sub-group comprising p times the number of known "substituters" should be formed and assigned a new I value (see section IV, 2) for substituters, I_s, equal to:

$$I_s = I \pm F\left[A - \left(m - \frac{r}{100} \cdot \frac{m}{12} \cdot W\right)\right] \qquad (16)$$

A, the mean post-partum amenorrhea value which was subtracted in calculating mean prolongation of stay, I, is restored for the substituting group value I_s on the assumption that current contraceptors are not amenorrheic, and certainly will not be when the assumed mean duration of use with the old method has elapsed.

For life-table calculations, a table for substituters is prepared in the normal way, but omitting allowance for amenorrhea. A separate negative summary, based on m months of use without attrition by discontinuation, but incorporating the estimated pregnancy rate with penalty, is prepared to yield the births that would have been averted by use of the first contraceptive method had substitution not occurred. This is deducted from net births averted. (See however, the conclusion of Section II.8(4).)

In choosing between the two principal methods available of performing the final calculations on the data, it would be invidious for the present author to pronounce. The summary calculations of prolongation of stay are mathematically neater and more sophisticated, but consequently less flexible and more fragile than the life-table approach.

Where the same parameters for durations per birth, proportions fertile at insertion, continuation and pregnancy rates are fed into the models, the results should differ little unless:

(i) continuation patterns depart radically from the exponential,
(ii) mean duration of use is long (over 5 years), or
(iii) a high proportion of users accepts near the end of reproductive life.

In these circumstances, the life-table approach seems the better adapted.

REFERENCES

Agarwala, S. N.

1966 "The arithmetic of sterilization in India", Eugenics Quarterly 13(3): 209.

Avery, R.

1973 "Patterns and implications of long term IUD use in Taichung, Taiwan", Studies in Family Planning, Vol. 4, No. 6: 139.

Barrett, J. C.

1969 "A Monte Carlo simulation model of human reproduction", Genus 25: 1.

1971 "Use of a fertility simulation model to refine measurement techniques", Demography 8: 481.

1972 "The time response in averted births", Population Studies, 26: 507.

Brass, W.

1970 "Outline of a simple birth distribution model for the study of systematic and chance components of variation", Proceedings of the Third Conference on the Mathematics of Population, Chicago.

Chandrasekaran, C. and M. Karkal

1972 "'Continuation rate', 'use effectiveness' and their assessment for the diaphragm and jelly method", Population Studies 26(3): 487.

Chandrasekaran, C., D. V. R. Murty and K. Srinivasan

1971 "Some problems in determining the number of acceptors needed in a family planning program to achieve a specified reduction in the birth rate", Population Studies 25(2): 303.

Chang, M. C., T. H. Liu and L. P. Chow

1969 "Study by matching of the demographic impact of an IUD program", Milbank Memorial Fund Quarterly 47: 137.

Chow, L. P.

1968 "Evaluation of the family planning program in Taiwan, Republic of China", Journal of the Formosan Medical Association 67: 187.

Freedman, R. and A. L. Adlakha

1968 "Recent fertility declines in Hong Kong: the role of a changing age structure", Population Studies 22: 181.

Freedman, R. and J. Y. Takeshita

1969 Family Planning in Taiwan: An Experiment in Social Change, Princeton, New Jersey. Princeton University Press.

Gopalaswami, R. A.

1962 "Family planning: outlook for government action in India", pp. 67-81 in C. V. Kiser (ed.), Research in Family Planning, Princeton, New Jersey, Princeton University Press.

Haynes, M. A., G. E. Immerwahr, A. George and P. S. J. Nayar

1969 "A study on the effectiveness of sterilizations in reducing the birth rate", Demography 6: 1.

Henry, L.

1953a Fécondité des Mariages. Nouvelle Méthode de Mesure, Paris, Presses Universitaires de France.

1953b "La fécondité des mariages au Japon", Population (Paris) 8:711.

1961 "La fécondité naturelle. Observation-théorie-résultats." Population (Paris), 16: 625.

Jain, A. K.

1969 "Relative fecundability of users and non-users of contraception", Social Biology 16: 39.

Kelly, W. J.

1971a "Estimation of contraceptive continuation functions", Demography 8: 335.

1971b "Estimation of births averted by family planning programs: the parity approach", Studies in Family Planning 2: 181, The Population Council, New York.

Keyfitz, N.

1971 "How birth control affects births", Social Biology 18(2): 109.

Lee, B. M. and J. Isbister

1966 "The impact of birth control programs on fertility", pp. 737-758 in Bernard Berelson *et al.* (eds.), Family Planning and Population Programs: A Review of World Developments, Chicago, University of Chicago Press.

Mauldin, P.

1968 "Births averted by family planning programs", Studies in Family Planning 33: 1.

Mauldin, P., D. Nortman and F. F. Stephan

1967 "Retention of IUDs: an international comparison", Studies in Family Planning 18: 1.

Mohapatra, P. S.

1966 "The effect of age at marriage and birth control practices on fertility differentials in Taiwan", unpublished Ph. D. dissertation, University of Michigan.

Mukerji, S. and K. Venkatacharya

1967 "Estimation of fecundability from a set of marital A.S.F.Rs and a study of the impact of a family planning programme on the natality of the population —a simulation model", pp. 912-922 in International Union for the Scientific Study of Population: Contributed Papers, Sydney Conference, 21-25 August.

Potter, R. G.

1963 "Birth intervals: structure and change", Population Studies, 17: 161.

1969a "Estimating births averted in a family planning program", pp. 413-434 in Behrman *et al.* (eds.), Fertility and Family Planning, Michigan, University of Michigan Press.

1969b "Renewal theory and births averted", prepared for the International Union for the Scientific Study of Population Conference, London, September.

1970 "Births averted by contraception: an approach through renewal theory", Theoretical Population Biology 1: 251.

1972a "Contraceptive impact over several generations", pp. 137-166 in: Population Dynamics, New York, Academic Press Inc.

1972b "Additional births averted when abortion is added to contraception", Studies in Family Planning 3: 53.

1972c "Births averted by induced abortion: an application of renewal theory", Theoretical Population Biology 3: 69.

Potter, R. G., R. Freedman and L. P. Chow

1968 "Taiwan's family planning program". Science 160: 848.

Potter, R. G., G. S. Masnick and M. Gendell

1973 "Postamenorrheic versus postpartum strategies of contraception", Demography 10: 99.

Potter, R. G., M. L. New, J. B. Wyon and J. E. Gordon

1965 "Applications of field studies to research on the physiology of human reproduction: lactation and its effects upon birth intervals in eleven Punjab villages, India." pp. 377-399 in Mindel C. Sheps and Jeanne Claire Ridley (eds.), Public Health and Population Change: Current Research Issues. Pittsburgh, University of Pittsburgh Press.

Ram, Bali

1971 "A demographic evaluation of the family planning programme in India", Journal of Family Welfare 18: 37.

Rele, J. R. and T. Patankar

1969 "Differential fertility of contraceptors and non-contraceptors", prepared for the International Union for the Scientific Study of Population Conference, London, September.

Ridley, J. C. and A. S. Clague

1972 "On births averted: assessment of three methods of estimation", prepared for the Conference on the Uses of Computer Simulation in Human Populations, June 12-14, Pennsylvania State University.

Ridley, J. C., M. C. Sheps, J. W. Lingner and J. A. Menken

1969 "On the apparent sub-fecundity of non-family planners", Social Biology 16: 24.

Ross, J. A.

1965 "Cost of family planning programs", pp. 759-778 in Bernard Berelson *et al.* (eds.), Family Planning and Population Programs: A Review of World Developments. Chicago, University of Chicago Press.

Scrimshaw, S. C.

1973 "Migration, urbanization and the family: a study among residents in the suburbios and tugurios of Guayaquil, Ecuador", unpublished paper prepared for the Division of Social and Administrative Sciences, International Institute for the Study of Human Reproduction, Columbia University, New York.

Seltzer, W.

1970 "Measurement of accomplishment: the evaluation of family planning efforts", Studies in Family Planning, No. 53 (May), pp. 9-16, The Population Council, New York.

Sharman, A.

1966 Reproductive Physiology of the Post-Partum Period. Edinburgh, E. and S. Livingstone Ltd.

Talwar, P. P., N. R. Parthasarathy and R. S. Kurup

1969 "On a method of calculating the number of births averted due to a family planning programme", E & I Section, Department of Family Planning, New Delhi, India.

Tietze, C. and S. Lewit

1968 "Statistical evaluation of contraceptive methods: use-effectiveness and extended use-effectiveness", Demography 5: 931.

Venkatacharya

1971a "A model to estimate births averted due to IUCDs and sterilizations", Demography 8: 491.

1971b "A Monte Carlo model for the study of human fertility under varying fecundability", Social Biology 18: 406.

1972 "The distribution of births averted over time due to various methods of family planning" In press.

Vincent, P.

1950 "La stérilité physiologique des populations", Population (Paris), 5: 45.

1956 Recherches sur la Fécondité. Paris, Presses Universitaires de France.

West Malaysia, National Family Planning Board

1968 Report on the West Malaysian Family Survey, 1966-67.

Westoff, C. F., R. G. Potter, P. C. Sagi and E. G. Mishler

1961 Family Growth in Metropolitan America. Princeton, New Jersey ,Princeton University Press.

Wolfers, D.

1968 "Determinants of birth intervals and their means", Population Studies 22: 253.

1969 "The demographic effects of a family planning programme", Population Studies 23: 111.

1970a Post-Partum Intra-Uterine Contraception in Singapore. D. Wolfers (ed.), Amsterdam, Excerpta Medica.

1970b "The Singapore family planning program: further evaluation data", American Journal of Public Health 60: 2354.

1971 "Contraceptive overlap with post-partum anovularity", Population Studies 25: 535.

1973 "Some problems in calculating births averted", paper prepared for the International Union for the Scientific Study of Population Conference, Liège, August 27-September 1.

Wolfers, D. and S. Scrimshaw

1974 "Child survival and interval between births in Guayaquil, Ecuador", In press.

Wolfers, D. and H. Wolfers

1968 "A brief report on family planning clinic clients four years after", unpublished.

Zatuchni, G. I. (ed.)

1970 Post-Partum Family Planning. New York, McGraw-Hill Book Co.

Chapter 7
MATCHING STUDIES

by
H. Bradley Wells

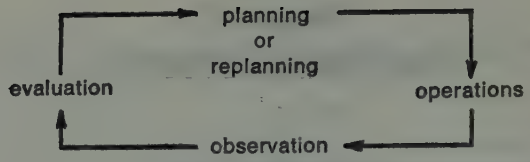
I. INTRODUCTION

Matching of variables is an accepted and useful method of adjusting for bias in observational (non-controlled) studies in which subjects self-select procedures or treatments, and can also be used to increase the efficiency of controlled trials (or experiments) in which subjects are randomly allocated to treatments. See Cochran (1965) for a detailed discussion of this. The distinction between the use of matching in the two situations is clear theoretically but is often ignored or blurred in actual practice in that matching in an observational study is often referred to as "controlling" for the intervening or disturbing variables. It would be more appropriate to refer to such studies as comparison rather than control studies.

In this chapter an attempt will be made to clarify the different purposes for which matching may be used, to discuss some of its limitations and to develop some recommendations for the use of matching in evaluating family planning programs. In order to place matching in perspective the development will proceed from the general to the particular, discussing first the program setting and evaluation, then briefly reviewing the principles of experimental design *vis-à-vis* observational studies as a background for discussing the use and limitations of matching and alternative approaches. Although the emphasis in this chapter is on the use of experimental and quasi-experimental trials and analytical procedures in program evaluation, it should be borne in mind that they comprise only part of the set of available objective and subjective techniques for use in an overall evaluation strategy.

II. THE PROGRAM SETTING AND EVALUATION

A family planning program can be viewed both as an *experiment* to achieve social goal(s) and as a dynamic evolutionary process with time and space dimensions, and which includes three major elements schematically related in a cyclic sequence as follows:



Improved societal well-being, the usual long-range objective of a program, is a very nebulous concept. More specific long-range objectives, such as improved maternal and infant health, reduced fertility levels, reduced rates of population growth, or prevention of unwanted births, are easier to evaluate. Fertility reduction is almost always an implicit, and often an explicit, long-range goal of family planning. Intermediate objectives would include increasing the level of use of effective contraceptive methods. Immediate objectives might also include raising the level of knowledge and attitude toward contraceptive use (Chandrasekaran and Freymann, 1965; Chandrasekaran, 1973). Operational objectives would include the provision of institutions, staff, and activities necessary to achieve the immediate and long-range objectives.

In this hierarchy, achievement of the immediate objectives is presumed essential for reaching intermediate objectives which in turn are essential for achieving the long-range or ultimate objective(s). Whether or not *the program* is the only way or even essential to achieving these goals has been questioned; but in this chapter these questions will be considered only in connection with the problem of allowing for non-program effects on achievement of program goals.

Program planning includes target setting at all levels and these must be time-oriented and must be revised as new experience is assimilated. A hierarchy of program operations designed to achieve the objectives in the target populations implies a timetable of activities and events which will take place in a prescribed time sequence to reach all areas of the state or country. In the absence of knowledge about the best way to achieve these objectives, experimentation with different operating procedures may well be part of such a program. Indeed it has recently been reiterated that controlled experimental approaches should be used more systematically in implementing social action programs (Light *et al.*, 1971).

Operations include the whole range of implementation activities from obtaining funds to staffing, training, setting up physical facilities, conducting education and service programs, and measuring services provided and their management. These operations have been termed "inputs" and their evaluation has been termed "quasi-evaluation", whereas "true evaluation" is based upon the degree of attainment

of immediate, intermediate and long-range goals (Greenberg, 1968). Development and testing of alternative operating approaches can be done on a sound basis if included as part of a well-designed and administered research effort.

Evaluation or assessment of the effect of a family planning program, as described in other chapters of this monograph, must be made in relation to the clearly defined goals and objectives of *the program.* This assessment is necessarily restricted to the program as it actually operates regardless of whether or not this actual operation corresponds at all with the planned operation. The evaluation process consists of a mixture of subjective and objective processes to determine the extent to which goals are being achieved and to provide data which can be used to plan for re-orientation of the program to achieve either previously specified or revised goals. See Caro (1971) for general references.

A critical examination should be made to determine the program operations or activities which could feasibly be subjected to controlled experimentation. This process will result in decisions about data collection and methods of analysis which would be appropriate for analyzing and evaluating the effect of those program components which cannot be subjected to controlled experimentation, and the end result would be a plan for the evaluation component in which each program element is classified by the manner in which it will be evaluated.

The methods of evaluation considered would include the whole range of procedures from controlled experimental trials (in which subjects can be assigned randomly to the procedure or program whose effect is to be determined), to methods of adjustment for potential biases, such as covariance analysis or matching studies (when subjects self-select procedures rather than being randomly assigned). Cochran (1965) described studies in which subjects cannot be randomly assigned to procedures as *observational* studies as contrasted with *controlled* studies. Many studies of cause-and-effect relationships in humans—for example, cigarettes and lung cancer—must of necessity be based upon observational rather than controlled studies, and are therefore subject to the criticism that other intervening or disturbing variables might be the cause(s). Estimating births averted or changes in fertility rates due to a family planning program are subject to similar limitations and criticisms. Cochran (1965) points out that

> "... an observational study is a study of the associations between two sets of variables ... and attempts to interpret these associations as causal or non-causal must rely heavily on information not supplied by the study."

In general, because of lack of randomization and replication, countrywide family planning, as well as other social action programs, fits the concept of a *natural* or uncontrolled trial more closely than the concept of a *controlled* trial. But it is possible, whether within the general hierarchy of objectives of a program or in preliminary trials or time phased implementation, to design and conduct controlled experiments for evaluating the relative worth of selected procedures and approaches. Even where controlled experimentation is impossible, it is useful for the evaluator to attempt to determine how a controlled experiment could be designed to answer the question (Dorn, 1953).

III. PRINCIPLES FOR DESIGN OF CONTROLLED TRIALS

Random assignment of experimental units or subjects to different procedures or treatments (programs), replication of trials, and objective measurement are the most essential features of controlled trials.

Methods for increasing the accuracy of experimental trials are discussed at length in books on experimental design; for example, Cochran and Cox (1957) includes a chapter devoted to this topic which includes consideration of:

(1) Increasing the size of the experiment (replication);
(2) Refinement of the experimental technique through:
    a) maintaining uniformity in the application of treatments,
    b) controlling the effect of extraneous factors (design and randomization),
    c) using unbiased measurement methods, and
    d) preventing gross errors;
(3) Reducing the effects of variability through:
    a) carefully selecting experimental units,
    b) taking additional measurements,
    c) grouping experimental units into more homogeneous groups.

Randomization is the most essential aspect of controlled experimental design. It provides added insurance for the experimentor in that objectivity is maintained in assigning experimental units of land, animals, or human subjects to the treatments or procedures to be evaluated. Randomization assures that potential bias, whether due to subject or experimentor, whether conscious or sub-conscious, does not influence the assignment of subjects to treatments and,

within the limits of probability, balances the influences of other potential disturbing variables by spreading them over all treatment groups. This permits the use of statistical methods for estimation and for the testing of hypotheses about differences between treatments. Control might include blocking, matching, or balancing of the potentially disturbing variables and it always includes random balancing of all other variables.

Neither unbiased nor precise measurements are necessarily assured by randomization. These depend both upon the degree of objectivity and the characteristics of the measurement technique. If the measurement process cannot be influenced by knowledge of the treatment to which the subject has been exposed, then it can be considered objective. The need for objective measurement has resulted in the so-called double-blind technique for evaluating the effects of drugs or other treatments on human or animal subjects in which neither the subject nor the scientist-evaluator measuring results knows which drug or treatment has been administered until the measurements have been made. Objective measurement techniques, however, are not necessarily unbiased; for example, interview methods in human populations are subject to many potential sources of bias from questions, respondents and field staff.

While replication, i.e., increasing the size of or the number of experiments, is essential, it does not assure representativeness unless the samples are randomly chosen from the target population about which inferences are to be made.

IV. DIFFICULTIES IN CONDUCTING CONTROLLED TRIALS IN SOCIAL ACTION PROGRAMS

One feature of experimentation in agriculture and industry as compared with human research is that the experimentor can "play God" i.e., the experimentor determines which plot of land or individual animal or batch of material will be assigned to a given treatment. For example, in testing the effect of different diets on growth rates of animals, no consideration need be given to individual animal diet preferences in assigning them to different diets, provided all other experimental design principles are followed. Ethical considerations dictate that testing the effect of different drugs or other "treatments" on human subjects must usually be restricted to volunteers (exceptions are public water supplies treated with chloride and/or fluoride but even in these "experiments" individuals are free to import their own water rather than use the public supply of treated water). Also,

it may be unethical to withhold a potentially beneficial treatment from a control group.

In social action programs strictly controlled experimental trials with random assignment of individual people to experimental or control treatments are seldom possible or feasible except with volunteers. For example, it is not possible to randomly assign women in the general population to noncontracepting or contracepting groups and even if they could be assigned in this fashion, it would be extremely difficult to be assured that they continue in the group to which they were assigned.

One can, however, use volunteers as subjects in certain experiments, e.g., women who want an IUD and agree to be subjects in an IUD experiment can be randomly assigned to different types or sizes of IUDs in order to determine the incidence of side effects and pregnancy (Tietze and Lewit, 1972). Similarly, different types of pills can be compared in a controlled trial. It would be impossible to conduct a strictly controlled trial of pills and IUDs in the double-blind sense with random assignment of volunteers to each group simply because they would know which treatment they were using and their response might be influenced by this knowledge.

To the extent that volunteers are representative of the "target" population, i.e., the population about which causal inferences are desired, then valid inferences can be made from planned experiments with volunteers. However, volunteers may have some other characteristic(s) associated both with the reason for volunteering and with the effects being tested so as to make them non-representative of the target population. Thus, generalization of results from self-selected volunteers to a larger population is fraught with hazards, a limitation that cannot be overcome even by controlled experimentation. But failure to use controlled experimental designs in testing treatments *among volunteers* may also lead to further biasing of results either because of patient or experimentor biases, or both.

Experimental designs may, however, be used with *groups* of people; for example, those residing in a given area may be used to test acceptance of social action programs where individuals are free either to join or not join the program, for example, to accept or not accept family planning services. The important distinction between the group versus individual approach is that the choice (acceptance or volunteering) comes afterwards and is, in fact, the response variable which the experiment is designed to measure.

An excellent example of the group design approach is the controlled study in Taichung in which the response variable was the acceptance rate of program contraception after random assignment

of small groups of people to three treatments in three strata (Berelson and Freedman, 1964). Even though cross-communication between experimental units occurred, valuable inferences were made as to which program efforts were more cost effective.

Differences in initial acceptance rates and relative costs for different treatment (program) approaches are very important for evaluating the operational, immediate and perhaps even intermediate program effects but they are of limited value in assessing the impact of different programs on fertility, the ultimate objective.

The all-pervading question, "What would the acceptors' fertility have been in the absence of the program?", must be answered or controlled for, in order to determine the effect of the program on fertility.

Greenberg (1968) discussed these difficulties and suggested approaches for applying sound but not strictly controlled experimental design methods for evaluating social programs in which areas or communities are used as experimental units. For convenience, Greenberg's groupings will be retained here:

(1) A major problem is establishing a set of comparable experimental areas (or subgroups within areas) which can be randomly assigned to different treatment procedures including serving as a control. Some communities will be ready and willing to implement a program but not all of these would be willing to participate as a control. Greenberg (1968) suggests testing alternative programs rather than program versus no program. Cross-communication between control and program areas must either be measured or kept to a minimum.

(2) Another difficulty is "confounding the merits of the program with the personnel involved" (Greenberg, 1968). For example, if program staff members prefer prescribing oral pills to intrauterine devices this will probably influence the choice of methods by individual patients even in a situation where free choice is available. Greenberg recommends rotation of staff to help in dealing with this problem. Staff members involved in measurement of program results might also bias results.

(3) The number and timing of measurements to be taken after the start of the program is another problem. Greenberg (1968) suggests several sample designs for taking repeated measurements in the different program areas in order to minimize the possibility that repeated measurements on the

same persons might influence program participation. For example, an initial KAP interview prior to introduction of a program might either influence acceptance of the program or it might influence response of an individual to a later follow-up interview. Cochran (1965) commented that "... an educational program that cannot improve health practices more than can a single questionnaire is not wrongly considered a failure ..."

(4) In a community program, if the double-blind principle cannot be used, one can at least try to "select observers who were not involved in providing service in that particular community" for measurement and evaluation (Greenberg, 1968).

(5) Greenberg (1968) feels that significance tests are probably wise "when in doubt about the appropriateness of doing so" but points out the need to stress that such testing is statistical as contrasted with the validation of social changes brought about by the program.

The Taichung study (Berelson and Freedman, 1964) illustrates the use of a controlled study using small areas within the city as experimental units to evaluate the extent to which women would volunteer for (accept) program contraception, an intermediate objective.

Greenberg *et al.* (1953) conducted a controlled experiment for evaluating the effectiveness of health education literature in changing nutritional and feeding practices of mothers with recently born infants. After mailing literature to the two groups of mothers, personal interviews were conducted to determine feeding practices. This study, in which sampling was based upon birth certificates, illustrates a technique which could be modified and used for evaluating either different motivational techniques in post-partum programs or any mailed literature approach to altering behavior, for example, acceptance of family planning.

Although there are serious difficulties, it seems fair to say that controlled experiments are not being utilized, perhaps rightfully so, to the extent that they might be in the evaluation of family planning programs.

V. HANDLING SOME DIFFICULTIES IN OBSERVATIONAL STUDIES OF SOCIAL ACTION PROGRAMS

Observational studies differ from controlled studies in that subjects or experimental units either have not or cannot be allocated

at random to treatments or procedures and often the conditions under which the study is carried out are not uniform either from place to place or time to time.

Because units or subjects are not assigned randomly to a treatment, observational studies are always open to the criticism that disturbing variable(s) other than the treatment might be responsible for the observed results. This criticism can only be combatted by repeatedly testing hypotheses about cause and effect in different situations, for example, see Hill's discussion of Cochran (1965) ([1]).

Matching, adjustment of subgroup means or rates and adjustment by regression (covariance) are related statistical techniques utilized to remove or adjust for potential biasing effects of major disturbing variables so that treatment effects can be more clearly seen. Even so these techniques do not necessarily remove all biases. Applications with a single disturbing variable are well illustrated in the literature (e.g. Cochran, 1953, 1965, 1968; Greenberg, 1953). Theoretically, extensions to multiple disturbing variables are relatively simple, but in practice it is difficult to satisfy the necessary conditions for valid application and the consequences are unknown. Even for a single disturbing variable the consequences of failure to meet the conditions for the various models have not been extensively investigated. Cochran (1965, 1968) and Rubin (1973a, 1973b) have examined some consequences by analytical methods. Billewiez (1965) and Rubin (1973a, 1973b) have used computer simulation methods to examine the statistical behavior and the relative merits of selected matching and regression adjustment procedures.

Cochran's (1965) paper and the ensuing discussion are highly recommended as a very practical guide to designing and analyzing observational studies. Rubin (1973a, 1973b) offers specific technical advice for the design of such studies.

In controlled experimental trials random assignment of units (subjects) to treatment groups ensures that sampling is from the *same population,* i.e., the initial treatment and control groups differ only by sampling error. Conversely, in observational studies the units (subjects) of their own volition "choose" one or the other group; hence, this in itself ensures that the two groups are from different populations. This usually means that the distribution of disturbing variables is different in the two populations. More than two groups can be compared, but for illustration it is sufficient to consider two with the understanding that one of the treatments can be either a lack of treatment ("control") or another treatment.

([1]) Hill's discussion follows Cochran's paper in the same issue of the journal.

The situation can be depicted generally as follows, using the notation for populations and sample size from Rubin (1973a, 1973b) and the lower case letters y and x to represent response and disturbing variables as in Cochran (1965).

Characteristic	*Population 1 (treatment 1)*	*Population 2 (treatment 2)*
Number of units	P_1	P_2
Random sample of units	$N = G_1$	$rN = G_2$
Matched sample of units	$N = G_1$	$N = G^*$
Response variable	y_1	y_2
Disturbing variables:		
First	x_{11}	x_{12}
Second	x_{21}	x_{22}
.		
.		
.		
j^{th}	x_{j1}	x_{j2}

Only one response variable, y, is considered but up to j disturbing variables, x's, are included although the methods and recommendations for continuous y and continuous x are based primarily upon analytical studies of y and a single x. Theoretically there is no limit to the number of variables which can be included in the adjustment process whether by matching, subgroup means or regression.

Practical considerations usually limit the number of x variables which can be used for matching whether they are continuous or discrete and with multiple x's it is difficult to determine whether the functional relationships are the same in both populations.

With regard to the populations, treatment 1 might, for example, represent program acceptors of family planning while treatment 2 might represent program non-acceptors. If it were desired to estimate the program impact on fertility rates or some other response variable, denoted by y, matched samples could be selected from the two populations, provided that P_2 is large enough. If desired, random samples may be taken, G_1 from P_1 and G_2 from P_2 and in order to match them, G_2 must be at least as large as G_1. If the distribution of disturbing variables is very different in the two populations or if exact matching is required G_2 must be several times as large as G_1. Rubin sets $G_2 = rN$, where $N = G_1$ and hence G_2 must be r times as

large as G_1, i.e., $r>1$. If no sampling is done, $G_1 = P_1$ or $G_2 = P_2$.

In his systematic review of problems in designing observational studies Cochran (1965) considers the major problems in detail under seven headings:

(1) setting up the comparisons (groups),
(2) the handling of disturbing variables,
(3) the step from association to causation,
(4) inferences from sample to population,
(5) measurement,
(6) multiple variables, and
(7) long-term studies.

Problems (1)-(4) are the major concern of this chapter because they are particularly difficult in observational studies. Problems (5)-(7), which are more general and present difficulties in controlled trials as well as observational studies, will not be considered further here.

In order to focus attention, the first four problems will be regrouped under three headings in the following sequence:

(1) Handling disturbing variables,
(2) Setting up comparison groups,
(3) Association and causal inference,

and will be discussed in relation to the question "How many births were averted by family planning program X?". More generally, the question might be posed as: "How many births would be averted in country Z in n years by program X?".

Estimating births averted by the program under a controlled trial could be a straightforward comparison between a number of randomly allocated program and control areas representative of country Z, following all the principles of sound design described in the sections above. Under these conditions major problems would be cross-communication, temporal changes and the number of years, n, over which the study would be conducted. However, if the family planning program were started in all areas more or less simultaneously without randomization (although not necessarily at the same level of efficiency in each case), an anomalous situation would result. Because there is no randomized control group, the evaluator is forced either to compare his area(s) with outside (pseudocontrol) group(s) or to resort to internal comparisons either between areas or between other subgroups exposed to his program. Greenberg (1968) suggests using

program subgroups with varying degrees of service as one alternative. Other alternatives are before-after comparisons with non-program areas or comparison of acceptors with non-acceptors.

It is informative to consider the design of a study to answer the question using Cochran's (1965) recommended approach. Decisions about response and disturbing variables must be made first because the variables to be included determine the types of comparison groups that should be chosen.

HANDLING DISTURBING VARIABLES

a) *Construct a list of known disturbing variables* arranged in three classes: (A) major variables for which it is essential to adjust, (B) secondary variables for which ideally it would be desirable to adjust, and (C) minor variables whose effects will be disregarded. This step requires a model or theory of the manner in which future fertility will be influenced by different variables, other than the program. One such list is shown below as a rough example, although it is recognized that the classification of variables according to their relative importance could well be different both for different cultures and for different investigators:

(A) *Major*	(B) *Secondary*	(C) *Minor*
Marital status	Age of wife at marriage	Value of children
Age of wife	Age of husband	Adult mortality levels
Number of living children by sex and age	Family structure	Future orientation
	Sterility	Religion
Parity	Fetal death history	Rate of modernization
Previous use of contraception	Infant mortality	Levels of fecundity
	Education of wife	Social norms about family size
Desire for more children	Occupation of husband	
Attitude toward family limitation	Income of family	Coital frequency
	Post-partum amenorrhea	
Use of abortion		

Although it serves for illustration, the list as it stands is obviously incomplete and the items lack the detail necessary to collect the required data even for those classified as major. A good source for expanding these lists and suggested detail regarding variables related to fertility is given by the United Nations (1970).

b) *Determine how the disturbing variables can be handled* to minimize their biasing effect and to maximize statistical efficiency. Consider first the variables in class B. Determine whether they are likely to bias comparisons in the response variables between treatment groups; if they are likely to bias results, shift them to class A, and if not, to class C. In pratice this step will not be possible if the

study is restricted to the use of existing records in which the disturbing variables are not recorded. Thus, additional data must usually be collected from samples of the two study groups in order to follow this good practice.

One condition necessary for the absence of bias in the comparison of two treated groups is that the frequency distribution of a single disturbing variable should be the same in the two groups; for multiple disturbing variables the multivariate distributions (not just the marginals) should be the same in both groups.

For both categorical and continuous variables the distributions should be examined. As pointed out by Cochran (1965), the usual practice is to use a chi-square test of the hypothesis of no difference between marginal distributions of single variables, and if the hypothesis is accepted, to proceed on the assumption that any bias due to the particular variable may be ignored. Two potential difficulties with this approach are:

(1) that equality of marginal distributions does not assure equality of multivariate distributions, and
(2) that the functional relationships between the disturbing and response variables(s) may be quite different in the two groups.

Billewicz (1968) draws attention to the fact that multiple disturbing variables are themselves often highly correlated and thus the number of matching criteria which can be "usefully employed" in efforts to minimize bias is limited.

If the mean of the disturbing variable(s) is equal in the two samples, this assures against bias due to linear regression; and if both the mean and variance of the disturbing variable(s) are equal, this assures against bias due to quadratic regression. As indicated by Cochran (1965) equality of higher moments assures against biasing effects of correspondingly higher degree polynomials. Since it is not clear what finding a non-significant difference in the distribution of disturbing variables implies, Cochran (1965) proposed an alternative approach to detecting possible biasing effects of the continuous disturbing variables in Class B, which he calls the "preliminary test of significance" of differences. He uses the following test criterion where x is the disturbing variable.

$$t = \sqrt{\frac{n}{2}}(\bar{x}_1 - \bar{x}_2)/\sigma_x \qquad (1)$$

where
n = the number in each sample,
$\bar{x}_1$ = mean for the sample from population 1,

$\bar{x}_2$ = mean for the sample from population 2,
σ_x = pooled standard deviation of x as estimated from available data.

This test can be based upon data available on the disturbing variables before the values of the response variable becomes known. Cochran states that if t is greater than 1.5 for a single disturbing variable, this suggests examining the relationship between that variable and the response variable after the values of the response variable are available to determine whether the class B disturbing variable should be shifted to class A or class C. If t is substantially above 1.5 for several disturbing variables of class B, the investigator must face the question of whether the groups are suitable for comparison since it raises doubts about whether the functional relationship between variables is the same in both groups.

Situations may arise in which estimates of mean differences between study groups may be desired even though functional relationships among variables differ in the two groups (Rubin, 1973a, b; Cochran, 1969).

Variables in class A, those which are definitely classified as disturbing variables, can be handled in a number of ways. Perhaps the most common are:

(1) *Pair matching*, which involves comparison and agreement of two individuals on all disturbing variables used in matching.
(2) *Mean matching*, also referred to as group matching or balancing, which involves cross tabulation of both groups by all disturbing variables and selection of the number from each cell corresponding to the number in the treated group.
(3) *Adjustment of subclass differences* which, as in mean matching, requires cross tabulation of both groups by all disturbing variables, but unlike mean matching no attempt is made to balance numbers; instead, means are calculated for each cell and a common set of weights is used to calculate an adjusted mean for each group.
(4) *Covariance adjustment*, which involves fitting a common (multiple) regression function to the disturbing variable(s) to determine the adjustment difference between the two groups.

Each method has advantages and limitations, both practical and theoretical. Both pair and mean matching require that the distributions of the disturbing variables in the two study populations (G_1 and G_2 on page 224) overlap sufficiently for the required matches to be

made. For example, if there are five categories of a disturbing variable in two populations and the distributions are as follows:

Number in population	Category 1	2	3	4	5	Total
P_1	0.10	0.20	0.30	0.30	0.10	1.00
P_2	0.20	0.30	0.30	0.15	0.05	1.00

a random sample from population 2 of at least 2 N, i.e., $G_2 = 2\,G_1$ would be required for mean matching a random sample of N from population 1. For pair matching or when the disturbing variable is continuous, a somewhat larger sample than size 2 N from population 2 would probably be required to match N pairs, depending upon the exactness of matching specifications. With multiple disturbing variables, the number of categories to be matched would be increased and it is quite likely that problems of matching would be increased unless the distribution of disturbing variables is very close in the two populations. Neither adjustment by subclass means nor by regression requires either equal sample sizes or equal numbers in the subclass and hence both methods may be applied to random samples from the two populations. However, as pointed out by Greenberg (1953) overlap of the distributions of the disturbing variable(s) is an implicit assumption in regression adjustment as it is for subclass means.

No assumptions about the functional relationship between response and disturbing variables are required except for the regression adjustment procedure, although it is questionable whether adjustment should be attempted at all in such cases unless the assumption is made that relationships are the same. Regression adjustment is usually restricted to use with continuous response and disturbing variables but it has also been applied in a simulation experiment with categorical disturbing variables and a continuous response variables (Billewicz, 1965). Recent developments also make possible the application of regression-like adjustments when response variable are categorical (Grizzle *et al.,* 1969), although the small sample properties of such estimates have not been analyzed even for the case in which both treatment and control samples are from the same population.

The variance of differences between treated and control samples and hence statistical efficiency differs for the four methods. When

the disturbing variable is categorical or discrete and measured without error, pair matching, mean matching and adjusted subclass measures remove all the bias, provided the relationship between the response and disturbing variables is the same in both populations (Cochran, 1965).

When the disturbing variable is continuous, pair matching removes any bias provided that the disturbing variable is correctly measured. Mean matching and adjusted subclass means may fail to remove all bias if the distribution of the disturbing variables in the two populations differs within the subclasses. When the relationship between the response variable and the disturbing variable is a linear regression, regression adjustment removes all the bias as does pair matching, whereas after mean matching or adjusting subclass means, slight bias remains.

However, if the relationship is other than linear regression, the first three methods would be likely to be superior to regression adjustment in removing bias. An advantage of regression analysis is that separate regression lines can be fitted and used to adjust data from the two samples, although as pointed out by Cochran (1965) and Greenberg (1953) one should be prepared to reconsider the question of whether the two groups should be compared if the regression slopes are different.

When the disturbing variable is subject to appreciable measurement error, none of the four methods completely removes bias (Cochran, 1965, 1968). Cochran (1965) summarizes as follows:

> "... similarities among the methods are greater than their differences. When feasible, matching is relatively effective. Overall, covariance seems relatively superior to adjustment by subclassification... If the original x (disturbing variable) distributions diverge widely, none of the methods can be trusted to remove all, or nearly all, of the bias."

Hence, the need to find comparison group(s) in which the distribution(s) of the disturbing variables are very similar.

SETTING UP COMPARISON GROUPS

To study cause and effect relationships in observational studies often involves selecting one or more comparison groups in which ideally "all other factors" than the presence or absence of the factor about which inferences are to be made are equal. The number of births averted by a program could be assessed by comparing areas with a program against those without, using either the total population or samples from the total population. Alternatively one could

reason that comparison of acceptors with non-acceptors of program contraceptives would provide a suitable estimate, *provided* all other relevant factors are closely matched and the functional relationships were the same. Since it is often difficult to determine whether "all other factors" are matched, the use of more than one comparison group is desirable. Often, only one of the groups, for example, program acceptors, can be definitely identified. Even then, their characteristics might vary with time and the investigator should take this into account, by comparing acceptors at different times with suitable comparison groups. Note that the characteristics of non-acceptors also change with time; for example, some non-acceptors who were pair matched with acceptors later became acceptors, thus eliminating the pair from the study (Takeshita *et al.,* 1964).

When the treatment (program) is applied to everyone then no comparison group can be found which was not exposed. This is the usual situation in evaluating acceptance of family planning, i.e., all women are exposed. The only recourse in this situation is to measure the response and disturbing variables at appropriate times before and after the program starts. If behavior is influenced by knowledge that the program is about to start it is desirable but not always feasible to measure the variables during a period prior to announcement of the program as well as after the announcement and after it starts.

Ideally the "before-after" designs should be applied to two groups randomly allocated to treatment and control, but as discussed elsewhere this is only possible with volunteers or when it is both ethically and politically feasible. Comparison of fertility levels of acceptors and non-acceptors before and after they accept is an attempt to simulate a "before-after" design. Similarly, areas may be "matched" on "all factors" other than the presence or absence of a program and measurements of trends compared over time (Rochat *et al.,* 1971).

In practice a decision to evaluate is often made so late that the only practicable approach is to collect data retrospectively. This is done by setting up groups that differ in the response variable and determining whether they differ in the suspected program variables, after adjustment for the possible effect of disturbing variables. Evaluating factors in acceptance through comparison of matched acceptors and non-acceptors would fall into this category. If the retrospective study is based upon the memory of respondents for causal factors in the distant past or on data from routine records, created for other purposes, the data may be of poor quality.

Although factors in acceptance and fertility rates are the two response variables used for illustration, the principles for setting up

comparison groups can be applied in evaluating a wide range of response variables in family planning programs.

ASSOCIATION AND CAUSAL INFERENCE

To make sound statistical inferences about cause and effect from observational studies about causal hypotheses in any social program, different evaluators may state the hypotheses differently, use different disturbing variables and different models of functional relationships, even at the same time or at different times within the same program. The picture is further confused by differences between different programs. This might be likened to the five blind men each describing the elephant while standing still and inspecting a different part of the anatomy.

Only by more systematic exploration of the whole phenomenon (or elephant) can more complete and consistent descriptions or hypotheses be developed. Standardization of measurement techniques and definitions helps in defining the phenomena to be measured and as in pure experimental designs the ability to replicate results is essential. Studies using the same design, however, may well be subject to the same biases and come up with the same or similar results; hence Cochran (1965) recommends using different designs which are not subject to the same biases, i.e., starting from a different vantage point on the elephant.

Stated more formally, "when constructing a causal hypothesis one should envisage as many *different* consequences of its truth as possible, and plan observational studies to discover whether each of these consequences is found to hold" (Cochran, 1965). For example, suppose the hypothesis is that the acceptance of program contraception, y, will increase as the level of program input, z, increases when all other factors, x, are taken into account, i.e., "are equal". If z is the amount of money *per capita* spent for the family planning program, use of three or more levels of z gives a more comprehensive check of the hypothesis than only two levels. If it is hypothesized that the relationship between y and z changes in specific direction(s) with changes in other variables (which have not been matched), say education, these relationships can be verified if the study is large enough. Other comparisons which should be made to test the hypothesis include acceptance rates,

(1) for the same amount of input over different lengths of time,
(2) for different amounts of input over the same period of time, and
(3) for areas in which input has been high and declined versus those in which it has steadily increased over time.

Moving to the next level we can hypothesize that acceptance of program contraception reduces fertility rates and the contrasts that can be set up to test it are more varied. Under the latter hypothesis, is the measure of acceptance rates.

In comparing acceptors with non-acceptors, z could be assigned the value of one for an acceptor and zero for a non-acceptor. If the convention is adopted that acceptance means that the acceptor will always be counted in that category regardless of later practice, it is clear that the size of a time cohort of acceptors in the program area can only decrease because of death, aging out of the fertile ages or migration out of the program area. Non-acceptor ranks, however, can be depleted by all of these factors as well as by becoming an acceptor. Thus, it is important to take these time factors into account in comparing "before-after" fertility trends of time cohorts of acceptors and non-acceptors at different points in time. Short-term differences between acceptors and non-acceptors at a specific time may disappear if followed for a sufficiently long time, *provided* that non-acceptors at that time are handled in the same way as acceptors, i.e., for later comparison with the acceptors at time t_0 all non-acceptors at t_0 are retained. Assuming a closed system, the situation can be depicted as follows:

acceptors	$a_0\ a_1\ a_2\ a_3\ a_4\ \ldots$
non-acceptors	$N_0\ N_1\ N_2\ N_3\ N_4\ \ldots$
time	$t_0\ t_1\ t_2\ t_3\ t_4\ \ldots$

The program starts at time t_0, a_1 accept in the interval (t_0, t_1) leaving N_1 non-acceptors at t_1 and so on. Thus, the future fertility of a_1 should always be compared with N_1 in later years regardless of whether some in N_1 later become acceptors. Similarly a_2 should be compared with N_2 etc. Thus, non-acceptors who accept will be counted in both groups but not in the same years since comparisons would always be time-specific.

If in addition the sample sizes were large enough to make age or marriage cohort specific comparisons, the causal hypothesis could be tested more rigorously. And in reporting the results of investigation(s) regarding the causal hypothesis one should always attempt to list and consider all alternative explanations of the results. This may require data from sources outside the study but if these alternative explanations can be anticipated in the planning stages, supplementary observations may be added to the study and materially improve it.

Table 7.1. *Comparison of selected features of the three studies*

Feature	*Study*		
	No. 1	*No. 2*	*No. 3*
1. Reference	Takeshita *et al.* (1964)	Chang *et al.* (1969)	Okada (1969)
2. Study population	Women in eight townships, Taiwan	Women in three counties, Taiwan	Women delivering in District of Columbia General Hospital, United States
3. Period of acceptance	April 1960-September 1961	January 1964-December 1966	November 1964-December 1965
4. Methods of birth control	Mainly condom and foam	IUDs	Mostly pills
5. Records matched with acceptor records	Household registers	Household registers	Birth certificates
6. Variables used for linking the two sets of records	Mother's name	Mother's name	Mother's name, hospital and date of delivery
7. Linkage difficulties reported	Migration before records linked	Migration before records linked. Address data on acceptor record incomplete	None
8. Records of acceptors not linked	Not reported	Twenty-six per cent estimated	Not reported
9. Variables in common record for matching to reduce bias	*a*) Age *b*) Marriage duration *c*) Number of living children	*a*) Age *b*) Education *c*) Interval since last live	*a*) Date of delivery *b*) Age *c*) Parity

	from the study		
11. Type of matching	Pair	Group (720)	Pair
12. Follow-up interview	No	No	Yes, acceptor and non-acceptors interviewed at least 18 months after delivery
13. Pairs eliminated after matching because of:	*a*) Acceptor pregnant within nine months *b*) Acceptor or non-acceptor moved out, divorced or widowed *c*) Non-acceptor becomes acceptor before cut-off date	Not applicable	Pairs in which one or more members not interviewed
14. Numbers of:			
a) Acceptor records	1,361	4,776	Not reported
b) Women in one source	1,361	4,300	680 (*a*)
c) Linked with other record	Not reported	3,181	680 (*a*)
d) Matched records	918	3,181	680
e) Ratio d÷b	0.67	0.74	Not comparable
f) Report based upon	918 pairs	3,181 pairs mean matched	200 pairs of whom 161 (80 per cent) pairs were interviewed
15. Acceptor vs. non-acceptor comparisons:			
a) Fertility prior to acceptance	5 years	3 years	Pregnancy rate prior to last delivery
b) Fertility after acceptance and/or births averted	6-30 months observation	18-42 months observation	12-month pregnancy rate after acceptance

(*a*) Records linked with acceptor record from a 20 per cent sample (1200) birth certificates for D.C. General Hospital.

VI. CRITIQUE OF SELECTED MATCHING STUDIES

The technique of age adjustment, a method familiar to demographers, is a specific form of adjustment by weighting of subgroup means, and adjustment by regression techniques has been considered elsewhere in this book; hence these methods will not be discussed further except as possible alternatives to matched sample study designs.

In this section the characteristics of three similar matching studies (Takeshita *et al.,* 1964; Chang *et al.,* 1969; Okada, 1969) will be compared in relation to their technical features (See table 7.1). The purpose of the comparison is to illustrate some of the questions and pitfalls which should be considered in designing other studies either to verify these findings or to extend them over time. One purpose of each of the studies was to compare pre- and post-acceptance differences in the fertility of acceptors with differences in fertility during the same period among non-acceptors. In each instance non-acceptors presumably have access to contraceptives from non-program sources. All three attempted to adjust for potential biases in estimates of fertility differences by matching acceptors and non-acceptors on disturbing variables although the procedures and specific matching variables differed somewhat between the studies. In both the Taiwan studies, the first stage required that records of acceptors be linked with household registers in order to create two study populations,

(1) acceptors for whom household registers are available,
(2) non-acceptors.

Matching on disturbing variables was done by comparing the characteristics of acceptors with those of non-acceptors to find suitable matches.

Okada (1969) on the other hand, studied women delivering in a single hospital who either accepted or did not accept post-partum program contraception within four months of delivery. However, she selected her study sample by first choosing a 20 per cent sample out of about 6,000 live birth records for the hospital and then linked these with the acceptor records to find 680 acceptors who were then matched with the birth records of non-acceptors ([2]). Okada also attempted to find and interview each member of each pair to obtain additional data, whereas no such attempt was made in either Taiwan

([2]) This implies that all matches outside the original sample of 1200 were also checked against the acceptor records in order to be sure that the match was not an acceptor. Presumably this procedure was followed because birth certificates are more readily accessible than hospital records.

study although Chang *et al.* (1969) indicated that it would be desirable to interview the non-acceptors to determine which contraceptive methods they were using.

Even though matching was done on certain important disturbing variables, all three studies are subject to the same inherent limitations:

(1) Acceptors and non-acceptors were not matched on their "motivation for family limitation" (Chang *et al.*, 1969). This could, of course, be included as part of a covariance design in future prospective studies—but no doubt would increase the costs quite a bit, since it would require measurement of this variable, presumably by interview (ideally at some time prior to the start of the program, but no later than the time of acceptance), for each acceptor and for a matching non-acceptor. It would also be desirable to subsequently follow up each woman and measure fertility, as well as changes in "motivation for family limitation", although fertility itself can presumably be measured through linkage with population or birth registers where they exist, as in Taiwan or West Malaysia ([3]).

(2) While all three studies showed large short-term differences in favor of lower fertility for acceptors, longer-range studies are required to determine the extent to which differences change or persist over the remaining reproductive life span. This is related to (1) above. Surely some non-acceptors as of time t_0 will later become acceptors. The analyst in longer term studies must decide which comparisons are most meaningful; see pages 232 and 233, for a more detailed discussion of this problem.

(3) A more subtle but potential danger somewhat analogous to the warning given by Greenberg (1968) is that different time trends in child-bearing or different cohorts may lead the investigator to make correct short-term conclusions, which do not hold for longer time periods.

Past fertility trends can be different and yet groups may match on cumulative fertility in some prior period. To illustrate, consider the three examples in Figure 7.1, which show three possible sets of two cohort age-specific fertility curves representing fertility of acceptors

([3]) Johnson *et al.* (1973) describe procedures followed for linking family planning acceptor records with birth records in West Malaysia through a seven-digit National Identification card number. This will also be used in an attempt to link non-acceptors to births for studies similar to the two studies in Taiwan.

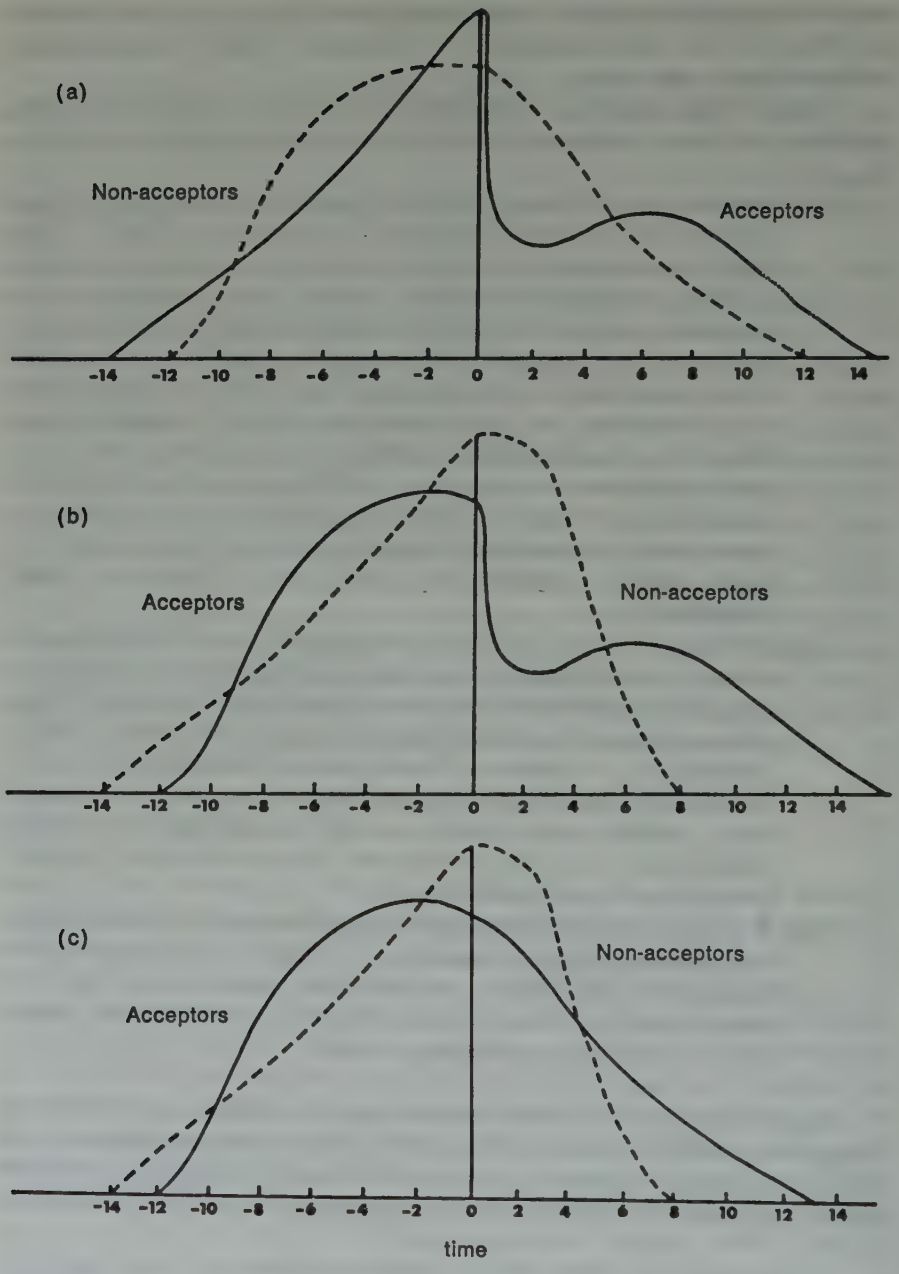

(a)
Non-acceptors
Acceptors
-14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14
(b)
Acceptors
Non-acceptors
-14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14
(c)
Non-acceptors
Acceptors
-14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14
time

Fig. 7.1. *Possible time (age) trends in cohort fertility.*

and non-acceptors at time zero; note that age can also be converted to correspond to the time scale in cohorts. The three main pairs of age patterns in Figures 7.1(a), 7.1(b) and 7.1(c) are the same up to time zero except that the dotted line representing non-acceptors in 7.1(a) corresponds to acceptors in 7.1(b) and 7.1(c), and vice versa. Areas under each curve up to time zero are equal assuring that age, parity and cumulated fertility rates are equal up to that point. Also, for a period of about three years preceding time zero the average fertility for the two cohorts is equal even though the two time trends are different. Further, the area under all six curves beyond time zero is equal, thus assuring that completed fertility is the same in all six cohorts and illustrating the point made in (2) above—namely that differences in fertility patterns might be reversed given time.

To speculate about the reality of the figures, let us consider them one at a time:

(1) If the pressure to use contraceptives is a fairly immediate consequence of child-bearing, Figure 7.1(a) might be a reasonable approximation depicting how those who experience a rapid spurt of child-bearing accept contraception, thereby causing their rates to drop rapidly, but then make up by increasing rates at a later time.

(2) If, however, there is a time lag before the pressure is sufficient to cause acceptance of contraception, Figure 7.1(b) might be more representative. Acceptors in this situation would not experience as great a drop in their rates after acceptance as those in Figure 7.1(a).

(3) If all acceptors are merely substituting program for other contraceptive methods, the situation might be represented as in Figure 7.1(c), with no abrupt change in the curve which had already started down at zero. Non-acceptors on the other hand begin to limit their births at a later time.

Note that this example has been highly simplified using only one age cohort and one time for illustration. It has also been assumed that all groups would have the same completed fertility.

In real life acceptors and non-acceptors are probably a mixture of an infinite number of much more varied patterns. Nevertheless, it might well be informative to attempt more refined cohort analyses based upon time (age) intervals shorter than one year. For example, Johnson *et al.* (1973) report that in West Malaysia 36 to 44 per cent of those accepting a contraceptive during the last month of a six-

month interval had a birth during that interval. This also suggests that seasonal fluctuations in births may well influence seasonal acceptance patterns.

Chang *et al.* used the open interval as one matching variable, but matching was not exact and dampened out the effects of variation in birth intervals since open intervals for non-acceptors were based upon time between date of last birth and July 1 of the year of acceptance. It is not clear what effect this might have on estimated differences in rates.

As described, the group of matching procedures of Chang *et al.* (1969) would have distributed the acceptors into 720 cells, within each of three calendar years. Similarly, non-acceptors were sorted into 720 cells, presumably for each year separately, although this is not clear from the description; so it is unclear what procedure was followed if matches could not be found. No data were shown for all non-acceptors. Acceptors in years two and three would have been non-acceptors in previous years. A matched non-acceptor " that had a live birth within nine months of July 1 of the year disqualified". Presumably a qualified substitute was added. No data were given on the number disqualified.

Takeshita *et al.* (1964) gave few details of the pair-matching procedure employed except that age matches were defined as agreement within five years. Okada (1969) listed the matching variables used but gave no details about the procedures employed or any description of matching tolerances.

Migration and loss-to-follow-up because of poor or incorrectly recorded data were mentioned as reasons for incomplete linkages of records from the two sources in both Taiwan studies but not by Okada. Matching biases may result from inaccuracies in the basic data.

A different type of matching study was done in Georgia (Rochat *et al.,* 1971), using seven pairs of small counties. One of each pair had an active family planning program for two or more years prior to January 1, 1968 and the other did not. The counties were matched on the basis of geographical location, racial distribution, and population size. Fertility trends were then compared before and after the program. This design is a partial approach to the recommendation that areas, or people, be grouped according to variations in the program (amount of services) and "before-after" comparisons made (Greenberg, 1968). As such the problems of coping with non-acceptors becoming acceptors is by-passed, but there can be cross-communication or spill-over effects; this is especially true of the pairs

of counties selected in this study since each pair was contiguous. Migration also makes it difficult to be very sure either of denominators or numerators.

Even though a paired design was used, Rochat *et al.* pooled all data for the seven program counties for comparison with the pooled results of the seven non-program counties, perhaps thereby obscuring some of the more precise comparisons that are theoretically possible with matched pairs designs. Since estimates were apparently available on the time of establishment of programs and utilization rates of family planning services in all counties, one might ask why they choose only 14 of the 159 counties in Georgia to study. Other designs using all counties might be more useful perhaps by using subgroup mean adjustment or by suitable regression analysis.

VII. SUMMARY AND DISCUSSION

Although the main focus of this chapter has been on the techniques of matching to reduce bias in observational as opposed to experimental studies, an attempt has been made to indicate, first, how observational and experimental studies may be used in evaluating programs and second, to make the point that matching is only one of several statistical approaches for handling biases due to disturbing variables. The other techniques include adjustment by weighted subclass means and by regression methods. Two types of matching designs may be used, pair matching and mean (group) matching. The subclass means technique is very similar to mean matching although it is simpler because the numbers in the subclasses are not required to be the same; thus, if data are already available for random samples from two populations, it will probably be simpler to use subgroup means than mean matching.

Pair matching is preferred if it is feasible. Factors of concern in pair matching are the number and accuracy of disturbing variables and whether or not available information might be unused because it does not match. Subgroup means and matching procedures have an advantage over regression adjustment in that specific functional relationships between response and disturbing variables need not be specified, although it is implicitly assumed that they are the same in both treatment comparison groups.

In designing observational studies to test causal hypotheses, one should always decide:

(1) how to handle all potentially disturbing variables,
(2) what comparison group should be studied to test the causal hypothesis, and
(3) to list all other possible consequences of the hypothesis so that, if possible, data can be collected to throw some light on these.

Review of four matching studies to estimate differences in fertility trends between acceptors and non-acceptors of program services was used to discuss a number of methodological issues. Unless "motivation to practice family planning" is added to future studies of fertility of acceptors and non-acceptors, it is doubtful that matching designs are suitable for providing much additional information about program effects. Analysis of data from studies like Chang *et al.* (1969) can be more efficient if more appropriate models which take account of short- and long-term trends in specific cohorts of acceptors and non-acceptors can be developed and used.

Ideally all of the above might be incorporated into studies for assessing births averted. The major addition would be the measurement of KAP levels at various times before and after the program is introduced. Another major change would be the creation of a number of comparison groups with different degrees of exposure to the program. If KAP determinations are known for a large number of women, for example from countrywide surveys, this information could be included either in matching or covariance analysis of subsequent acceptance or non-acceptance as well as fertility performance. Ideally the group, or sub-group, with known KAP levels prior to the program should be followed individually and important events related to fertility measured and recorded at prescribed time intervals, say each year. This would include repeated measurement of KAP for specific cohorts of acceptors and non-acceptors, including strength of desire for family limitation. It would then be possible to form comparison groups with different patterns of prior use and exposure to the program for analysis. The aim would be to account for all groups in the population by degree of exposure to the program and to determine the number of births occurring in each. No account has been taken here of the cost as well as measurement and logistical problems implicit in these recommendations. Studies of fertility which include the dynamics of "desire for family limitation" will without doubt be more costly than studies which do not; but if better evidence on births averted by family planning programs is needed, this expense must be borne.

REFERENCES

Berelson, B. and R. Freedman

1964 "A Study in Fertility Control", Scientific American: 29-37.

Billewicz, W. Z.

1965 "The Efficiency of Matched Samples: An Empirical Investigation", Biometrics, 21: 623-643.

Caro, Francis G.

1971 Readings in Evaluation Research, Russel Sage Foundation, New York.

Chandrasekaran, C.

1973 "A General Frame for Evaluation of Family Planning Programs", Paper presented at the Annual Meeting of the Population Association of America, New Orleans, Louisiana, April.

Chandrasekaran, C. and M. W. Freymann

1965 "Evaluating Community Family Planning Programs" in M. C. Sheps and J. C. Ridley, (Editors), Public Health and Population Change, University of Pittsburg Press: 266-286.

Chang, M. C., T. H. Lin and L. P. Chow

1969 "Study by Matching of the Demographic Impact of an IUD Program", Milbank Memorial Fund Quarterly 58: 137-157.

Cochran, W. G.

1953 "Matching in Analytic Studies", American Journal of Public Health, 43: 684-691.

1965 "The Planning of Observational Studies of Human Populations", Journal of the Royal Statistical Society A, 128: 234-265.

1968 "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies", Biometrics, 24: 295-313.

1969 "The Use of Covariance in Observational Studies", Applied Statistics, 18: 270-275.

Cochran, W. G. and G. M. Cox

1957 Experimental Designs, 2nd. edition, John Wiley and Sons Inc.

Dorn, H. F.

1953 "Philosophy of Inferences from Retrospective Studies", American Journal of Public Health, 43, 677-683.

Greenberg, B. G.

1953 "The Use of Analysis of Covariance and Balancing in Analytical Surveys", American Journal of Public Health 43: 692-699.

1968 "Evaluation of Social Programs", Review of the International Statistical Institute 36: 260-278.

Greenberg, B. G., M. E. Harris, C. F. MacKinnon and S. S. Chipman

1953 "A Method for Evaluating the Effectiveness of Health Education Literature", American Journal of Public Health 43: 1147-1155.

Grizzle, J. E., C. F. Starmer and G. G. Koch

1969 "Analysis of Categorical Data by Linear Models", Biometrics 25: 489-504.

Johnson, J. T., T. B. Ann and L. Corsa

1973 "Assessment of Family Planning Programme Effects on Births: Preliminary Results Obtained through Direct Matching of Birth and Programme Acceptor Records", Population Studies 27: 85-96.

Light, R. L., F. Mosteller and H. S. Winokur Jr.

1971 "Using Controlled Field Studies to Improve Public Policy", Chapter 6 in Report of the President's Commission on Federal Statistics, U.S. Government Printing Office.

Okada, L. M.

1969 "Use of Matched Pairs in Evaluation of a Birth Control Program", Public Health Reports 84: 445-450.

Rochat, R. W., C. W. Tyler Jr. and A. K. Shoenbucher

1971 "The Effects of Family Planning in Georgia on Fertility in Selected Rural Counties" in Advances in Planned Parenthood, 6, Proceedings of the American Association of Planned Parenthood Physicians, Boston, 9-10 April, 1970: Excerpta Medica Foundation.

Rubin, D. B.

1973a "The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies", Biometrics 29: 185-203, March.

1973b "Matching to Remove Bias in Observational Studies", Biometrics 29: 159-183, March.

Takeshita, J. Y., J. Y. Peng and P. K. C. Liu

1964 "A Study of the Effectiveness of the Prepregnancy Health Program in Taiwan", Eugenics Quarterly 11, December.

Tietze, C. and S. Lewit

1972 "Comparison of the Copper "T" and Loop "D": A Research Report", Studies in Family Planning, 3: 277.

United Nations

1970 Variables and Questionnaires for Comparative Fertility Surveys, ST/SOA/SER. A/45.

ACKNOWLEDGEMENTS

This work was partially supported by Research Grant HDO3341 from the U.S. National Institute of Child Health and Human Development.

Chapter 8
REGRESSION ANALYSIS OF AREAL DATA

by
Albert I. Hermalin

I. INTRODUCTION

The use of areas, rather than individuals, as the unit of analysis has long been a prominent feature of demographic research. Its role in such research has been influenced by the very nature of demography, in which the concept of population as an aggregate of individuals is central, by the type of demographic data available, and by the nature of the problem. In studying the factors associated with differentials in fertility, a cursory historical review since 1880 suggests a trend away from reliance on areal units towards emphasis on data specific to individuals (or families). In the earlier period, the lack of data on individuals constrained researchers of differential fertility to rely heavily on areal comparisons (see Lorimer, 1959, pp. 144-145 on this and for references to many of the studies). Emphasis on characteristics of individuals grew with recognition of the many social, economic and psychological factors that could affect fertility as it became more and more a matter of individual decision-making, and as research into such factors became more feasible with the development of the sample survey as a reliable procedure for the collection of data. General studies of differential fertility with areas as the unit of analysis have persisted, however, as exemplified in the work of Collver *et al.* (1967), Heer and Boynton (1970) and Nerlove and Schultz (1970). Nevertheless, in recent years areal analysis has been most often employed in studying the relation of fertility to economic development, with whole countries the usual unit, and in historical demography, with subnational areas the typical unit. As an example of the former see Friedlander and Silver (1967) which cites many other works of this genre and for the latter, the European fertility project as described by Coale (1969) and exemplified for Portugal by Livi-Bacci (1971). Areal analysis has also been employed

in evaluating the effectiveness of family planning programs and it is this application which forms the basis of the present chapter, though much of the discussion will pertain equally well to the other applications. Lastly, note should be taken of studies which combine aspects of both the individual and areal modes such as those by Duncan (1964) and Rhodes (1971). In these studies the aim is to appraise the relative importance of individual characteristics vis-à-vis those of the environment—as reflected in areal values—as determinants of individual fertility.

The aim of this review is to discuss the rationale for the use of areal analysis in family planning evaluation; to describe and illustrate the types of data employed and the methods of analysis; to point up the cautions and limitations which must be observed; and to suggest some directions for improving this method of research.

II. STRUCTURE AND RATIONALE OF MULTIVARIATE AREAL ANALYSIS

Areal analysis has been used in a number of ways in attempting to appraise the effect of a family planning program. Areal units have been the basis of experimental designs; fertility rates in areas with and without programs (or with more and less intensive programs) have been compared, occasionally with efforts to match the areas on other factors relevant to fertility; and the timing of fertility change in relation to a program in one or more areas has been compared with the rate of change in non-program areas. Studies utilizing areas in these manners have been reviewed and summarized by Ross *et al.* (1972, 3-5). A conceptually distinct approach, and the one on which this paper focuses, is to use characteristics of areal units as the basis of a multivariate analysis.([1]) This method of evaluating a family planning program typically has the following features:

([1]) The term "multivariate analysis" has come to have both a general and a specific meaning. In general usage, it refers to the situation where one has two or more variables for a set of units. In its more specific sense, multivariate analysis refers to a situation where one has two or more dependent variables. In the latter case the term multiple regression (or another term depending on the specific technique) would characterize the situation of one dependent variable and two or more independent variables. In this paper we use "multivariate analysis" in the more general sense, and will present examples of both a single dependent variable and two or more dependent variables.

(i) for all or some subset of local areas within a country or other territorial unit, one has a series of characteristics whose influence on fertility is to be determined (these are the independent variables);
(ii) one or more of these measures reflects inputs into the program, either in terms of manpower, supplies, money, or some intermediate effect of such inputs such as acceptances;
(iii) a measure of fertility (or some characteristics assumed to bear directly on fertility) as dependent variable, which reflects the effects of the program and other factors; and
(iv) a method of analysis which enables one to appraise the relative impact of the independent variables on the dependent variable.

This sketch of a multivariate areal analysis suggests that its structure is particularly congruent with the objectives of a family planning program and the form in which questions concerning the effectiveness of such a program have been posed. A primary objective of family planning programs is to reduce aggregate fertility levels by increasing the availability of family limitation methods. Thus, the use of an aggregate fertility measure as dependent variable is consonant with this aim. In addition, countries with family planning programs are also undergoing socio-economic change and the question arises as to the extent to which decreasing levels of fertility are due to increasing levels of education, decreasing child mortality and other modernizing trends, rather than the program itself (Davis, 1967). A reasonable test of this contention is to see whether among areas within a country, there is a relationship between levels of fertility and program inputs or acceptance rates, after taking into account socio-economic and demographic factors.

Viewed this way, it should be clear that areal analysis may be the preferred mode of testing a hypothesis and not merely a second-best alternative when individual correlations are unavailable.

The interrelations between an analysis based on areas as aggregates of individuals and one based on individuals must be clearly understood to avoid errors of inference and to use either or both modes to maximum advantage. The most well known problem associated with areal analysis is the so-called "ecological fallacy" as described by Robinson (1950). The point here is that correlations observed to hold among areas cannot be assumed to hold among individuals, so that from an areal analysis which shows, for example, lower fertility in areas of higher program acceptance, one cannot be certain of finding that acceptors have fewer children when a correlation based on individuals is carried out. The statistical interrelations

of correlations based on areas and those based on individuals are set forth by Duncan *et al.* (1961, pp. 65-67) who also describe various sets of conditions which, if assumed to hold, allow inferences about relationships on the individual level on the basis of relationships on the areal level. (See also Hannan, 1971, 38-49; and Alker, 1969). Typically the investigator will have little or no evidence as to the tenability of those assumptions, so that considerable risk attaches to disaggregating areal results. Insofar as the problem has been posed as one properly to be investigated at the areal level, and inferences remain at that level, this danger is avoided. The problem of statistical inference based on aggregate data is also treated in the econometrics literature, generally under the heading of aggregation bias. (See Theil, 1954, for an analysis of the problem from this standpoint.)

The possibility of change in the correlation or regression coefficient from one level of analysis to another suggests that attention be given to the source and nature of the conceptual models employed at each level. A micro-theory of fertility behavior which is fruitful for explaining variation among individuals may not satisfactorily account for areal variation. In addition, the ranking of the importance of the explanatory variables may be different in the two cases. By the same token, a satisfactory analysis of areal variation may not identify important variables for explaining individual behavior. Because of this, any model which claims to explain both individual and areal variation should be tested at each level since success at one level does not imply success at the other (Duncan *et al.*, 1961, pp. 100-102).

Aside from the statistical behavior of similar variables at the individual and areal levels, models focusing on one or another level are likely to employ different concepts. An analysis of fertility at the areal level might include such characteristics as differential access to family planning knowledge and services, differential effectiveness of family planning inputs, factors of transportation, terrain, etc. Some of these areal characteristics may not have any precise counterpart on the individual level, being functions of the level of aggregation, while others represent factors which tend to be overlooked and/or be difficult to measure when conceptualization takes place at the individual level. Further distinction among types of areal variables is given in Section V.

The foregoing problems do not preclude fruitful interaction between approaches. The conceptual model for an areal analysis may be strengthened by studying micro-theories of individual fertility behavior; results from an areal analysis may suggest hypotheses and variables that should be pursued on an individual level; and as noted

in the Introduction, where the data permit, both individual and areal level variables can be employed in a single analysis.

In this section we have sought to make clear the basic structure of an areal multivariate analysis employed in family planning evaluation and its general relation to an evaluation based on individuals as units of analysis. We turn next to several studies which illustrate the use of this approach, returning in a later section to examine several specific conceptual and statistical problems associated with areal multivariate analysis.

III. EXAMPLES OF AREAL ANALYSIS IN PROGRAM EVALUATION (2)

Table 8.1 outlines the elements of a number of multivariate areal analyses which bear on family planning program evaluation. We will describe each of these briefly in order to elucidate the concepts employed, the hypotheses or models tested, the techniques utilized, and the data that were available and unavailable to the investigators. This review will then form the basis of some summarizing ideas for these components of the procedure.

Simmons' (1971; 110) objective was to determine if there was a causal relation between the government family planning program in India and the acceptance of IUDs and sterilization. He reasoned that if there was such a relationship, "we would expect to find a statistically significant association between acceptance and variables measuring the program inputs. Alternatively, if it was the level of development that explained acceptance, then one would expect that socio-economic variables explain most of the variation in acceptance rates and that policy variables are statistically insignificant." To test this he employed a linear multiple regression in which acceptors of an IUD or sterilization per 1,000 total population were regressed on a measure of program input and three measures of socio-economic development. The unit of analysis was the district and these data were available for 246 districts.(3) The measure of program input used was the number of clinics per population, which Simmons regarded as a weak measure since it does not fully capture the quantity of program effort nor reflect at all the quality of the family planning organization. The socio-economic characteristics employed

(2) Many of the technical terms and concepts reviewed will be treated in detail in subsequent sections.

(3) For analyses which use the states of India as units, see Agarwala (1972) and Vig (1972).

Table 8.1. *Elements of several multiv*

Country	*Dependent variable*	*Indepe* *Demographic-ecological*	*Socio-economic*
India	Acceptors of IUD or sterilization per 1,000 pop.	% urban	% literate Per capita income
Korea	I. Couple-years-of-protection—total and by type of cp.	Children 0-4/women Women 20-44/pop.	Radios/household Non-farm pop./pop
	II. Cost per CYP	Children 0-4/women Density	Same as I plus: High school grad/p
Taiwan	I. Age/sex standardized C.B.R. Age specific fertility rates	Reciprocal of child survival probability	Proportion of 12-19 school Proportion of those and not in schoo primary school education Proportion of male force in agricult
	II. Age specific fertility rates Age specific percent married IUD acceptors/ women 20-44 Doctors/women 15-49	Same as I plus: Age-standardized C.B.R. (a) Percent married (a) Aboriginal or non-aboriginal area Taichung city or not Density	Primary school gra pop. 12+ (for females and tota Per capita income Male/female educ level
Taiwan	Age specific marital fertility Total fertility rate Age specific percent married	Child mortality Age at marriage	Proportion of male force in agricult Percent females 1 in school Primary school gra females 12+ no school

(a) Endogenous variable.

nalyses related to program evaluation

ogram input	*Number and type of areas*	*Years covered*	*Multivariate technique*	*Investigator and source*
pop.	246 districts	1966	Linear multiple regression	Simmons (1971)
s/pop. /pop. ibution of CYP pe of cp. ors/women	185 health center areas (coteriminous with county and city boundaries)	1969	Linear multiple regression (with 2nd order terms)	Riew (1971)
months/ en 15-49	361 administrative areas	1964-1968	Linear multiple regression (with 2nd and 3rd order terms) Log linear model Non-linear estimation First difference model	Schultz (1971a)
s I plus: ipients per en 15-49 for s 1964 ceptors/ en 20-44 (a) / en 15-49 (a)		1964 1966 1968 (for fertility)	Simultaneous equation model (Two stage least squares)	Schultz (1971b)
ecific IUD otance of married en	331 administrative areas	1966-1969	Linear multiple regression (path analysis)	Hermalin (1971)

were proportions of the district's population who are literate, the proportion residing in urban communities (usually defined as 5,000 or more in size) and *per capita* income (only a 10-year-old measure was available). The partial correlation coefficient of acceptances with the input measure was substantially larger than the coefficients obtained with the three socioeconomic characteristics used in the multiple regression equation.

A subsidiary analysis of 56 districts for which a second input variable—the ratio of extension workers to population—was available, revealed that both input variables were strongly related to the acceptance rate, while controlling for the socioeconomic characteristics. The partial regression coefficients were also used to calculate the partial elasticities which show the percentage change in the dependent variable associated with a one percent increase in each independent variable.

Simmons found the analysis "largely consistent with the hypothesis that past variation in the number of family planning acceptors by district has been the result of variations in the inputs to the Family Planning Program... The implication is that if the situation prevailing in 1966 continues, the more inputs to the Family Planning Program, the greater will be the output." (Simmons, 1971; 119.)

In the Korean study, based on 184 health center areas, Riew (1971; 4) attempted "to identify various socioeconomic factors associated with the achievement and the cost of the family planning program". The analysis involved two sets of regressions. In the first set the dependent variables were measures of program achievement while in the second they were cost estimates per unit of program result.

The explanatory variables shown in Table 8.1 were used in a set of linear multiple regressions with estimates of couples-years-of-protection (CYP) for each type of contraception (IUDs, vasectomies, condoms and oral contraceptives) and with a weighted CYP from all forms as the dependent variable. The independent variable measuring field worker inputs was used both in its linear and quadratic form in the equations.

The results, taking CYP from all forms of contraception as the dependent variable, indicated highly significant positive partial regression coefficients with the proportion urban and worker inputs (in linear form) and a strong negative effect from the proportion of physicians. The other independent variables were positively associated but not significantly so. The pattern of effects also varied with the type of contraception studied. Riew interpreted the effect of proportion urban as reflecting both the heightened efficiency of the

program due to the greater density and the higher education in urban areas. The negative relation with the number of private physicians suggests that on balance they were competitive with the program, the increase in availability of private family planning services which they represented more than off-setting the services they performed for the public program.

From the coefficients of the linear and quadratic family planning worker variable, Riew was able to determine the level at which adding workers would no longer increase achievement.

In the second set of equations, Riew used various combinations of the explanatory variables shown in Table 8.1 with cost per couple-year-of-protection as the dependent variable. In general he found (p. 28) that cost decreased with the scale of the operation (measured in terms of CYP) and education, increased with the availability of radios (assumed to serve as a proxy for income), and was strongly associated with the mix of contraceptive services.

Schultz's work in Taiwan was based on a more explicit conceptual model of fertility determination and used a wider variety of statistical techniques than that of other investigators.(4) The model was based on considerations of the environmental constraints and opportunities that influence the number of surviving children parents want, the level of child mortality, and the family planning activities viewed as facilitating the avoidance of unwanted births. The model was avowedly individualistic in formulation and the results are often so interpreted, raising a number of the considerations discussed in the previous section of this paper. Data for areas were used in the absence of comparable data for individual families.

In the first series of regressions, the three environmental variables used were:

(i) education of children 12-19 (reflecting the cost of child rearing);
(ii) the extent of adult education (assumed indicative of a demand for higher quality children and ability to avert unwanted births); and
(iii) the proportion of the male labor force in agriculture (under the assumption that a child's labor is more valuable in an agricultural setting).

Separate measures of worker input were used for the two types of workers employed by the program, and second and third order terms of these inputs were introduced to allow determination of returns to scale. The environmental and mortality variables were

(4) In addition to the models described in Table 8.1, see Schultz (1973) for other specifications of the determinants of fertility in Taiwan.

lagged three years and the input variables one year to simulate the delays involved in their impact on current fertility.

The results of the analysis suggest that, aside from program input, child mortality and children's education are the two major influences on fertility. Schultz found that the worker inputs had a negative impact on fertility in the early years of the program but that there were sharply diminishing returns with increases in personnel, so that these variables were no longer significant at the end of the period. His results also indicated differential effectiveness for the two types of workers employed by the program, suggesting that overall impact would have been enhanced by a different mix of these personnel.

Schultz's second model employs a simultaneous equation system to jointly determine fertility, proportions married, IUD acceptance rates, and the doctor to women ratio. These variables were each determined by a number of exogenous variables shown in Table 8.1 and by some of the endogenous factors. Thus, the doctor to women ratio was determined by a number of exogenous variables; the IUD acceptance rate by the doctor variable as well as exogenous variables; the proportion married by the standardized crude birth rate as endogenous variable plus exogenous variables; and the fertility measures by proportion married, IUD acceptance rates, and doctor to women ratio, treated endogenously, as well as exogenous variables. To avoid bias in the estimators, two stage least square estimation procedures were used, with ordinary least squares regression results presented for comparative purposes.

From the equations determining age-specific fertility, child mortality (positively) and child education (negatively) appeared to have strong influences. The other socioeconomic characteristics varied more from age group to age group; the IUD acceptance rate did not appear significant for most age groups and the worker-months of family planning effort displayed a generally weak negative association which varied considerably among age groups and among the years studied. Substantial differences between estimators obtained by two stage least squares and ordinary least squares were indicated.

The conceptual approach used by Hermalin in his analysis of the Taiwan program differs from that employed by Schultz. His model postulates that an area's level of socioeconomic development, its demographic status in terms of age at marriage and child mortality, and its level of IUD acceptance from the program, influence its level of fertility. Hermalin utilized the three socioeconomic characteristics shown in Table 8.1 as indicators of the influence of modernization and urbanization on fertility rather than as measures of the importance

of the specific characteristic in question. Other possible socio-economic characteristics were excluded to avoid the problem of multi-collinearity, which makes it difficult to reliably gauge the individual effects of highly intercorrelated variables (Johnston, 1963; 201-207) and the problem of redundancy arising from the tendency of a cluster of intercorrelated variables to divide up their joint influence and thereby reduce the apparent influence of any one characteristic (Gordon, 1968). Though Hermalin's approach utilizes a macro-model orientation, which may be more appropriate to areal data, it sacrifices much of the explicitness apparent in Schultz's parental decision-making model.

The analysis utilizes a recursive path-analytic model. Age-specific fertility, for the four five-year age groups between 20 and 40, is treated as the product of proportion married and age-specific marital fertility at each age.([5]) The logarithm of this identity produces an additive equation for these two components of age-specific fertility. Each of those components is then regressed on a different combination of the variables shown in Table 8.1. Age at marriage and the IUD acceptance rates are treated as endogenous variables but ordinary least squares regression is considered applicable, under the assumptions used.

Data are utilized for 331 of the 361 administrative areas of Taiwan, the 30 aboriginal areas being excluded due to their distinctness on a number of factors. In addition, the 331 areas are subdivided into three groups according to the level of marital fertility prior to the program and the regression analysis is carried out separately for each group. The results indicate that at the younger ages, age at marriage and the IUD acceptance rate (in the higher fertility group of areas) are significant influences on marital fertility, while at the older ages, child mortality, child education and IUD acceptance rate are the important influences. There is a particularly strong negative relation between marital fertility and IUD acceptance rates at older ages in the subgroup with high preprogram fertility levels, this subgroup also being the least modern. This suggests that the program has been successful in these areas and that acceptance of program contraception in the more modern areas may to a large extent represent substitution of methods.([6])

([5]) Both legitimate and illegitimate births—the latter a small factor—are included in the numerator of age-specific marital fertility.

([6]) These results are suggestive rather than providing a rigorous test of this hypothesis because dividing the sample according to the level of the dependent variable may introduce bias into the estimation of the parameters. A similar result obtained by Schultz (1973) using an alternative approach is described briefly in section VIII.

As a partial test of the effect of differences in estimating techniques, two stage least squares was applied to a simplified model and the results were closely comparable to those obtained from ordinary regression procedures.

This review of the work of four investigators shows that despite the same general framework, there can be considerable difference at each stage of the process. The outcome of an analysis is often sensitive to variation at each of these stages. Thus, investigators using the same model and data, but employing different statistical techniques may obtain quite different results. Similarly, differences may arise from alternate operational definitions of the concepts, although the other elements remain fixed. For this reason, we examine each of these stages more closely in the following sections, pointing out its importance for the overall procedure, the options available, the limitations and problems associated with this stage and possible improvements. For the purposes of this discussion the elements of the areal multivariate procedure will be grouped into the following five categories:

1. Purpose of the analyses; conceptual framework; specification of the model;
2. Data utilized; operationalization of concepts; types of data typically available and their sources; data not often available; validity of data;
3. Choice of areal units;
4. Multivariate techniques; choice of techniques; results obtained from each technique;
5. Interpretation and use of results.

IV. CONCEPTUAL FRAMEWORK AND SPECIFICATION

The analyses described in the previous section vary considerably in their conceptual frameworks. At one extreme, Riew's work may be deemed heuristic, in that it seeks to establish the nature and magnitude of the relation of various factors to fertility-relevant variables, but the work is not motivated by any prior formulation of the basic relationships. As such, any hypotheses derived from the analysis cannot be taken as confirmed by the data employed, though the empirical relationships discovered can serve as inputs for future model building. At the other extreme, Schultz's approach starts with a well-formulated, albeit micro-level, conceptual framework which implies a set of structural relations, and his interest is in estimating the magnitude of these relationships.

An intermediate position of conceptual explicitness characterizes ne work of Simmons and Hermalin. Though Simmons' expressed ntent is to test a hypothesis regarding program inputs and accept- nces, the hypothesis is rather general and necessitates a specifica- on of other factors affecting the dependent variable. These other actors are set forth only in general terms as socio-economic charac- eristics or level of development, and no further rationale is provided or choosing one or another specific variable. Hermalin presents a omewhat more explicit model than Simmons but again the variables hosen to operationalize the level of modernization or development re arrived at in an *ad hoc* fashion.

Any attempt to explain, interpret, or synthesize an empirical set f relationships requires a theory, model or conceptual framework. We use these terms interchangeably here although further distinction s possible.) A theory is an arbitrary construction which identifies he key elements underlying a set of phenomena, it is more or less xplicit, and more than one theory may be advanced for a given ubject matter. As such, theories are best regarded not as true or alse, but as more or less fruitful. The criteria usually employed for ppraising theories are their ability to account for established rela- ionships and to discover new relationships and, somewhat second- rily, parsimony. Theories are abandoned or modified when they annot account for newly observed phenomena or when observations ontrary to their implications occur. Theories may also be modified y extending a parsimonious list of key concepts to include variables hat more fully explicate the mechanisms or processes which account or the observed relations. Duncan *et al.* (1968a; 16) have termed his approach "an incremental strategy of model building".

In recent years, interest in the determinants of fertility and family lanning has led to the development of a number of conceptual chemes of a social, psychological, or economic nature (Freedman, 967; Fawcett, 1970; Chung *et al.*, 1972; Schultz, Theodore W. (ed)., 973). In the main, these schemes have been intra- rather than nterdisciplinary; many are general frameworks for research which equire further delimitation before yielding a set of structural relations vhich explicate a set of phenomena. Though some of these frame- vorks are designed to explain inter-societal differences in fertility, nost of them focus on individual differences. In addition, few of he available models fully explicate how a family planning program ffects the fertility process. Such aspects as the way in which pro- ram activity may lead couples to substitute program contraceptives or non-program ones, as well as encouraging current non-users to dopt contraceptives from both program and non-program sources,

should be incorporated into a model which attempts to evalute the effects of a program. Conceptualization of these processes can aid in the choice of the program characteristics (e.g., field worker activity versus total expenditures) to be incorporated in the analysis.

The importance of a theoretical formulation is highlighted when attention is given to the statistical models employed to test hypotheses or estimate parameters of the structural relations. The theory identifies the key concepts and their interrelationships. It is these concepts which must be operationalized in order to generate the data to be used in testing and estimation.

The assumptions of the theory and of the data generated—the scale of measurement, the nature of the distributions, the form of the relationships, the nature of unmeasured influences and the like—are used to identify the mathematical model isomorphic with this structure. Application of this mathematical model then becomes the basis for manipulating the data so as to permit statistical inference and the estimation of unknown parameters.

As an example, take the well-known case of a single equation regression model which may be written as:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \mu_i \qquad (1)$$

where,

Y_i represents fertility of the i^{th} area;
X_{1i} is a measure of educational level;
X_{2i} is a measure of infant mortality;
X_{3i} is the rate of acceptance of program contraception;
μ_i is the error or disturbance term representing all unmeasured influences on fertility.

Use of this mathematical model for statistical inference entails a long list of assumptions, including:

(i) the data are interval scales (or dichotomous); ([7])
(ii) fertility is a linear function of the other characteristics;
(iii) there is no reciprocal influence from fertility to the other variables;
(iv) the disturbance term populations each have zero mean and equal variance;
(v) the error terms are uncorrelated with one another;
(vi) the error term is not correlated with any of the predictor variables.

If these assumptions are met by the theory and data, then least square procedures can be used to obtain estimates of the population

([7]) The standard regression model, however, is not applicable when the dependent variable is dichotomous (see Nerlove and Press, 1973).

parameters (in this case the β's of the equation and the variance of the disturbance populations). These estimates contribute to our understanding of fertility levels by revealing the relative importance of the factors affecting fertility and since these factors can be influenced to a greater or lesser degree by social measures, the results have implications for public policy. To the extent that the assumptions are not met, then estimates from the use of the mathematical model may be biased and/or misleading.

Of the assumptions listed, we focus at this point on the requirement that the error term not be correlated with the predictor variables. Let us assume that age at marriage affects fertility and is correlated with educational level and acceptance rates. A model which incorporates age at marriage explicitly into the equation will produce different estimates for the effect of education and acceptances than one which leaves this as part of the unmeasured influence, thereby altering our understanding of the fertility process and perhaps the implications for program policy. Failure to include a variable of this sort is referred to as specification error. Safeguards against this problem require that careful attention be given to all the factors thought to be important determinants of fertility and to the nature of their relationships. This is not to suggest that one include in an equation every factor having a possible relation to fertility. A catalog of such factors is not a theory and, as will be pointed out, may itself produce misleading inferences. Goldberg (1966) has suggested that the uncorrelated error term assumption is "almost always invalid" and that efforts to identify a salient element of the error term, when observations contrary to a theory's prediction occur, is one source of theory modification.

Very often the explication of the conceptual model will involve a number of equations. The second model of Schultz described in Section III is one such example. Here recognition is made of the detailed interrelationships among the variables. Thus, the theory relating marriage and fertility in this model leads to treating the proportions married and the fertility rates at each age as simultaneously and jointly determined. The simultaneity arises because in this formulation "the probability that a woman of a specific age is currently married has a bearing on her desired and actual fertility. Conversely, her reproductive goals influence the probability that she will be married at any specific point in time." (Schultz, 1971b; 8.)

Many models of fertility, particularly those that emphasize the role of household decision-making, can be expressed adequately only through a system of structural equations. Where this is the case, it is important that the full system be developed, even though current

interest focuses on only one. This is so because the appropriate estimation technique for even a single equation which is part of a system must take into account the nature of the assumptions about the correlation of error terms across equations. Quite obviously, this must also be done when the focus is on the complete system and one wishes to estimate all the parameters (Johnston, 1963; 231).

In Section VII, the explication of a model in terms of a series of equations, and the conditions under which various estimating techniques may be used, are taken up in some detail.

The aim of this section has been to stress the importance of the theoretical stage in undertaking areal multivariate analysis for appraising the impact of family planning programs. Models for explaining areal variation have not received the same degree of explicit attention as those dealing with individual differences. The adequacy of any theory rests upon its ability to account for observed relationships and on the congruence of its assumptions with the mathematical models underlying the estimation procedures.

V. SOURCES AND NATURE OF DATA

Empirical analysis cannot begin until operational definitions (also called rules of correspondence, and empirical definitions, among other terms) are established for the concepts of the theory. Application of the operational procedure generates the data for the investigation. A useful distinction can be made between unobservable and observable concepts. In the latter, the investigator has or can devise an operational definition which, it is felt, adequately measures or captures the concept. Thus, a procedure for recording deaths among live births can develop a metric which adequately reflects the concept "infant mortality". Unobservable concepts are more highly abstract and measurements related to these concepts would be regarded as indicators of the theoretical terms rather than operational definitions (Blalock, 1963). For example, measurements representing level of education, occupational structure, and mass-media exposure might be regarded as indicators of the underlying concept "modernization". Though the distinction is one of degree, it is nevertheless important. Developments in causal model analysis can provide inferences about the causes and effects of unobservables in properly specified models, thereby producing closer concordance between theoretical reasoning and analysis of observable data. (These techniques will be discussed in the next section.)

Unobservable concepts may also arise from the lack of data

and/or the inability to execute a desired operational procedure. Inferences about such unmeasured variables are also possible under appropriate circumstances. Since unobservables of both kinds are likely to arise in the areal analysis of program effects, these possibilities enrich the potential contribution of this mode of evaluation. Whether a concept is regarded as observable or not is in part a function of the theoretical model. It may be noted that areal data measurements may more often be properly regarded as indicators where a micro-model of fertility determination is employed than when a macro-theory is utilized. For example, the educational level of the children in an area may be regarded as an indicator of the concept "parents' aspirations for children's education" or an operational definition of the macro-concept "educational opportunities".

As the last example suggests, a given concept may be operationalized in different ways and the same empirical data may be used as operational definitions for a number of different concepts. The studies previously reviewed illustrate these elements of arbitrariness. Simmons uses the proportion literate in each area as "an important index of social development" and as "a measure of receptiveness to different sources of information" (Simmons, 1971; 112). Schultz uses a measure of adult education because the extent of such education "may stimulate demand for higher cost (quality) children and also lower the subjective and pecuniary cost of averting 'unwanted' births" (Schultz, 1971a; 19). Hermalin uses an educational variable similar to that of Schultz as a representative of the influence of modernization and urbanization on fertility. As another example, Simmons' measure of family planning program input is number of clinics; Riew uses family planning workers per 10,000 *persons*; Schultz employs worker-months per 1,000 *women* 15-49; and Hermalin worker-months per 10,000 *married women* aged 20-44; Riew uses radios per 100 households as an explanatory variable and expects its relationship to contraceptive use to be positive because this is the most common mass medium, but he conjectures that the variable may also be serving as a proxy for income.

While recognizing the data constraints and differential objectives faced by the investigators, it is worth noting Duncan's *et al.* (1968a; 44) admonition that inadequate "attention to standardization and replication of research procedures" impedes cumulative research and the comparability of findings.

Keeping in mind that the data to be sought are dictated by the concepts of the theoretical model, we review below the types of variables likely to be available for an areal analysis of program effects, their sources, and desirable data not often available. The outline is

not meant to typify the situation in any one country though it is obviously influenced by the analyses already carried out.

Generally an evaluation of program effects using areal data will require data of a demographic, socioeconomic, and family planning program nature. Specific variables that may be available by area are as follows:

A. *DEMOGRAPHIC AND ECOLOGICAL VARIABLES*

1. Fertility: crude birth rate; general fertility rate; age-specific fertility rates; total fertility rate; child-women rates.
2. Nuptiality: marital status by sex and age or proportion married by age; marriages, number and/or rates; marriages by age.
3. Migration and mobility rates; migration streams.
4. Mortality: crude death rate; infant mortality rate; other age-specific mortality rates.
5. Population: total number; sex and/or age distribution; sex ratio.
6. Density: population potential; distances to urban centers; level of urbanization; land use measures.

B. *SOCIOECONOMIC AND RELATED VARIABLES*

1. Economic activity status (labor force)—by sex, age; occupational distribution (detailed or agricultural vs. non-agricultural); other job characteristics.
2. Adult educational level (per cent literate; detailed educational attainment—by sex, age).
3. Ethnic, religious characteristics.
4. School attendance measures—primary, secondary.
5. Income measures—*per capita*, family or household.
6. Housing characteristics; ownership of specified objects.
7. Commercial activity; manufacturing and agricultural production; number of establishments; tax revenues.
8. Communication and transportation measures (postal service volume; number of telephones, etc.).
9. Health and welfare measures: number of physicians, hospitals, clinics, etc.; health budgets; health services—maternal and child care (proportion of births in hospital, with doctor in attendance), number of examinations for specific diseases; incidence and prevalence rates; social service budgets; public work expenditures.
10. Civic measures such as proportion of adults voting.

C. *FAMILY PLANNING PROGRAM CHARACTERISTICS*

1. Program inputs: number of field workers and other personnel; number of clinics and other facilities; physicians and other medical personnel employed or under contract; man-hours of efforts; number of contacts; expenditures; promotional efforts (number of posters, screenings, etc.).
2. Number and characteristics of acceptors of contraceptive service; type of contraception accepted; first contacts vs. subsequent contacts.

The above list is not meant to be exhaustive. Few countries will have data on all or most of these characteristics available for areal subdivisions and many will have data for only a small number. In most cases fertility will be a key dependent variable in the model, in which case the availability of a suitable measure is a precondition for undertaking the analysis. Quite obviously, the crude birth rate with its dependence on the age structure is less satisfactory than some of the other measures but given age and sex population data, standardized measures can be developed. Note should be taken, however, that a set of areal indirectly standardized birth rates does not strictly control for age-sex compositional differences (Kitagawa, 1964). Beyond the fertility variables, the analyst will have to investigate whether the data called for by the conceptual model are available, keeping in mind that some compromises may be necessary.

The areal variables in the above list are in Lazarsfeld's (1958; 111-114) phrase, "properties of collectives", and using his classification one may distinguish among them those which are analytical properties—formed "from some mathematical operation upon some property of each single member"—and global properties—"those not based on information about the properties of individual members". [These do not exhaust Lazarsfeld's classification. Further useful distinctions between measures of aggregates and those of individuals may be found in Kendall and Lazarsfeld (1955; 290-296) and Duncan *et al.* (1961; 41-46).] This distinction and the nature of some of the variables outlined above highlight the point made earlier, that areal analysis need not be simply an individual analysis "writ large".

To obtain the available data one will have to consult a number of statistical sources of the country of interest. Census reports, vital statistics publications, and/or data from a household registration system can be expected to contain information on some of the variables enumerated in categories A and B above. Variables in category C will generally come from the service statistics of the agency responsible for the family planning program. For Taiwan, much of the

available data can be obtained from two series of publications—the annual *Taiwan Demographic Fact Book* (Taiwan Provincial Department of Civil Affairs) and the annual *Family Planning Reference Book* (Taiwan Provincial Institute of Family Planning). In addition, general statistical compendiums for a country may contain data on some of the socioeconomic variables in category B, and one may also consult the reports of specific departments and ministries, such as transportation, health, education, agriculture, manpower-development and the like. Occasionally the local governing unit will publish data specific for that area and, where the statistics are comparable, compilation of these may provide data not otherwise available. In addition to data which are routinely collected and published, there may be special studies conducted by one or more agencies which will prove useful.

For any multivariate analysis the variables to be employed must be available for all or substantially all of the designated units. Consideration must be given therefore to whether the variables of interest relate to the same administrative divisions and subdivisions. Where there is a well-established classification of geographic areas this is not likely to be a serious problem, but even in this case one may find special areal definitions such as health districts or school districts whose boundaries differ from those applicable to other variables. In an analysis covering a series of years, one should also check for changes in the classification scheme and changes in the boundaries of individual areas. One probable source of difficulty is the likelihood that some variables of interest will be available only for primary geographic divisions while others can be obtained for secondary or tertiary subdivisions. Though the detailed data may be aggregated into the broader unit, the analyst may prefer to carry out the investigation with the larger number of smaller areal units. Countries establishing family planning programs might benefit by checking that the resulting service statistics by area coincide in definition with statistical series already established.

Limitations may also arise from the frequency with which certain data appear. Decennial census data alone invoke obvious constraints on the timing of family planning evaluations. Nerlove and Schultz (1970; 36-37) used interpolated census data along with yearly birth and death statistics, and discuss the conditions under which such a procedure seems warranted.

One type of data rarely accessible on an areal level is information on attitudes, expectations or intentions. Such variables as desired or expected family size, wanted and unwanted pregnancies, attitudes toward contraception, aspirations for children's education and expectations concerning support from children often figure prominently in

models of fertility determination. Data on these subjects are not usually included in a census or other sources of areal statistics. One possibility for capturing such data resides in the KAP and other large-scale surveys frequently undertaken in countries with family planning programs. Sampling for these surveys usually involves, as a first step, a selection of secondary or tertiary geographic subdivisions of the country. If the number of areas in the sample and the sample size per area are sufficiently large, then individual data can be aggregated to form areal characteristics not otherwise available. These characteristics can then be combined with the more "standard" areal variables for a subsidiary analysis of these areal units.

Data on contraceptive behavior outside the family planning program is another important topic for which areal data are often unavailable. Such data can be strategic in helping determine whether acceptance from the program represents an addition to or substitution for existing contraceptive practice. One may distinguish two different approaches. In the first case, data from acceptors on use or non-use of contraception prior to accepting a method from the program can be developed from the program's service statistics. This requires obtaining this information from each recipient of contraceptive services and aggregating the data on an areal basis. Though not frequently done, development of this information is not inherently difficult, and its widespread adoption would add a potentially useful variable for areal analysis. A second approach is again through the use of KAP surveys. As before, the contraceptive information collected in these surveys can be aggregated, when the number of cases and number of areas warrant, to form areal statistics on the extent of contraceptive use outside the program.

Several countries may have available additional areal detail from the operation of the family planning program which can prove highly useful in improving the operation of the program, even if not employed in appraising the program's impact on fertility. Data on the characteristics of the workers in each area (age, marital status, education, etc.) and the quantity and type of field work, such as the number of contacts of each type—first visits, revisits, group meetings, etc.—can be combined with characteristics of the area in appraising the relative importance of each factor in the number of acceptors recruited. The analysis of this type for Taiwan by Chen and Hermalin (1970) suggested, for example, that married women rather than unmarried be recruited for field workers, that workers be left to accumulate experience in one area rather than be moved around, and other policy program implications.

In contrast with the situation in conducting a sample survey, the

investigator employing an areal analysis typically will not have any control over the procedures involved in collecting, tabulating, and preparing the information utilized in the study. For the most part, the data used are published statistics from primary or secondary sources, and this lack of control may lead to a lack of concern about the level of measurement error. This question merits attention, since measurement error can significantly affect the results of the analysis. This is seen most directly in the case of a simple regression analysis: if there is random measurement error in the independent variable, the regression coefficient will be biased and, for large samples, will be underestimated as a function of the measurement-error variance relative to the true-value variance (Johnston, 1963; 148-150). Developments in the analysis of measurement error also permit one to trace the implications of various patterns of non-random or correlated measurement error, which can arise when the measurement error is correlated with the variable being measured or where there is correlation among the measurement errors in two or more variables (Siegel and Hodge, 1968; Chai, 1971; Schoenberg, 1972a). By incorporating alternative formulations of measurement error with the posited causal models it may be possible, with appropriate assumptions, to evaluate the relative plausibility of different models in the light of the observed data. Valuable as these approaches are in making explicit the assumed interrelationships among variables and measurement errors, they do not remove the desirability of reducing measurement error to the greatest extent possible. As Schoenberg (1972a; 31) notes, "measurement error must eventually be eliminated at the source because the data-analysis methods for accounting for the presence of measurement error are ultimately only simplifications of the reality and consequently only approximately cover it up." In the case of an areal analysis, the investigator may be well advised to consider the following steps:

(1) To review any available information on the completeness and reliability of the data utilized. For censuses and vital statistics, many countries conduct evaluations of the census in accord with United Nations recommendations, and one may consult the United Nations *Demographic Yearbooks* for a broad appraisal of the vital statistics situation in a particular country. The investigator may also use well-known demographic techniques to carry out analytical tests of data adequacy.

(2) For other data, familiarization with procedures used to collect and process the information may prove helpful. For example, knowing the definitions of the events in question, regulations concerning the

degree of local authority, and the extent to which the data are tabulated centrally, afford some clue as to variation which may arise from those aspects. Knowledge of the definitions and procedures is also useful, of course, in deciding whether a given statistical series is appropriate as an indicator or operational definition of a theoretical construct.

(3) Where data are available for each area on a regular basis, it is possible to carry out consistency checks to help detect such things as typographical errors and transpositions of data between areas. One simple procedure is to compute the annual percentage change in an item for each area and to examine the instances of large deviations from the national average, taking into account the nature of the item and the size of the area. (An area with a small population, for example, might show large percentage changes in the crude birth rate.) Where both the detail and total are shown for a characteristic, a check on the addition can easily be carried out; rates and other measures can be recomputed when the appropriate "raw" data are given. Errors found in this way can often be corrected from internal evidence; in other instances it may be necessary to inquire of the responsible agency. The experience of this writer is that even data of generally high quality may contain a large number of such errors and that many of them can be detected by these simple procedures. It should be noted that in multivariate areal analysis one is often dealing with a relatively small number of units, so that large errors in a few observations can have a sizable effect on the correlation and regression coefficients.

(4) In carrying out the analysis, assumptions about measurement errors should be incorporated into the model so that their effects are reflected in deriving estimates of the parameters.

Knowledge of general socioeconomic conditions of the country in question may sometimes prove helpful in choosing appropriate data for areal analysis. As Schultz (1971a; 22) has noted in a somewhat different context, interareal variation in school attendance rates may be largely due to areal differences in the supply of schooling or in the demand for schooling on the part of parents. If the variation is due mainly to differential capacities of the school systems, attendance data may be a poor operational definition of a theoretical construct dealing with parental interest in their children's education.

Many of the points discussed in this and the preceding section can be illustrated by the following analysis, in which a comparison is made

(i) of the results obtained by two analysts using the same conceptual model and the same operational definitions;
(ii) of the effect of varying the operational definitions slightly; and
(iii) of the effect of varying the conceptual model slightly.

Since both this writer and Schultz have utilized the same published data in their analyses of Taiwan, we present first the results obtained in replicating Schultz's linear regression model (Schultz, 1971a; 49). The definitions of the variables and their notation as reported by Schultz are shown in Table 8.2. The dependent variables

Table 8.2. *Definition of variables used in Schultz linear regression model*

B_t	:	Birth rate in year t, normalized by national birth rate in specific year and standardized for age/sex composition in the case of the crude birth rate.
D_t	:	Reciprocal of the child survival probability from birth to age 15 in year t.
S_t	:	Proportion of persons enrolled in school over age 11 of the number of children 12 to 19 in year t. Observations in 1961 and 1965 are used to linearly interpolate estimates for intervening years.
E_t	:	Proportion of adults over age 12 who are no longer in school who have completed a primary education in year t. Observations in 1961 and 1965 are used to interpolate linearly estimates for intervening years.
A_t	:	Proportion of male labor force employed in agriculture in year t. Observations in 1961 and 1965 are used to interpolate linearly estimates for intervening years.
V_{t-1}	:	Number of man-months of Village Health Education Nurses' time allocated to the family planning program through year t—1 per thousand women in the region between the ages of 15 and 49.
P_{t-1}	:	Number of man-months of Pre-Pregnancy Health Workers' time allocated to the family planning program through year t—1 per thousand women in the region between the ages of 15 and 49.

Source: Schultz, 1971a; 53-55.

in the model are the standardized and normalized crude birth rate and the normalized age-specific birth rates. The child mortality, education and labor force measures are lagged three years, and the field worker input measures are lagged one year. The procedure followed was to use the basic data set utilized by Hermalin to construct the variables, precisely as defined in Table 8.2, and then to carry out the specified linear regressions. The only point of ambiguity arose in the definition of the cumulative worker-input rates, since Schultz

Table 8.3. *Correlations between fertility, environmental and program input variables used in Schultz linear regression model for Taiwan, compared with independent replication*

	Schultz	*Schultz replication* (a)	*Hermalin input measures* (b)
B_{1968} (C.B.R.) $\times D_{1965}$	0.692	0.712	
B_{1968} (C.B.R.) $\times S_{1965}$	−0.517	−0.528	
B_{1968} (C.B.R.) $\times E_{1965}$	0.234	−0.163	
B_{1968} (C.B.R.) $\times A_{1965}$	0.367	0.510	
$D_{1965} \times S_{1965}$	−0.373	−0.320	
$D_{1965} \times E_{1965}$	0.375	0.044	
$D_{1965} \times A_{1965}$	0.303	0.368	
$S_{1965} \times E_{1965}$	−0.073	0.452	
$S_{1965} \times A_{1965}$	−0.554	−0.695	
$E_{1965} \times A_{1965}$	0.192	−0.512	
$M_{1967} \times B_{1968}$ (C.B.R.)	−0.042	−0.041	−0.056
$M_{1967} \times D_{1965}$	−0.077	−0.093	−0.107
$M_{1967} \times S_{1965}$	−0.104	−0.055	−0.042
$M_{1967} \times E_{1965}$	−0.049	−0.100	−0.093
$M_{1967} \times A_{1965}$	0.039	0.014	0.011
$F_{1967} \times B_{1968}$ (C.B.R.)	−0.105	−0.140	−0.164
$F_{1967} \times D_{1965}$	−0.199	−0.214	−0.228
$F_{1967} \times S_{1965}$	−0.040	−0.001	0.020
$F_{1967} \times E_{1965}$	−0.084	−0.116	−0.106
$F_{1967} \times A_{1965}$	0.016	0.008	0.009

(*a*) Worker inputs represent cumulative worker-months of effort through 1967 per 1,000 women 15-49 as of June 1965.

(*b*) Worker inputs represent cumulative worker-months of effort through 1967 per 1,000 married women 20-44.

Notes: See Table 8.2 for definition of variables. Correlations are based on 361 administrative areas.

Sources: Column 1 - Schultz (1971a; Table H-4, p. 76).
Columns 2 and 3 - Tabulations performed at Population Studies Center, University of Michigan.

does not indicate the point in time used for the denominator of the cumulative inputs or whether the rates themselves were cumulated over time.(8) In the replication presented here, the cumulative inputs

(8) A private communication from T. Paul Schultz indicates that the yearly rates were cumulated. The population mid-year base was obtained for select years and estimates for the intermediate years were obtained by interpolation. The differences in means and standard deviation between the original Schultz measures and the replication are fairly small, as indicated by a comparison of the first two columns of the supplement to Table 8.4.

through 1967 were divided by the number of women as of mid-year 1965, roughly mid-way through the operation of the program as of that time.

Before presenting the comparison, we wish to stress that there is no implication in this analysis that one result is more correct than another. The point at issue is to illustrate how much difference can arise when two analysts using the same data-set process these data independently. Differences may arise from differentials in the recording and punching of the data, from differentials in the consistency checks undertaken and the nature of the corrections made, from differences in the procedures used to construct the variables employed, and possibly from the computer programs utilized. Such comparisons are all too rare in the social sciences, and it is salutary to study the effect of differential processing as well as differentials arising from alternate conceptual models and estimation procedures.

To avoid the ambiguity possible from incomplete knowledge of how the worker-input rates were defined, the top panel of Table 8.3 compares the correlations among the other independent variables and their correlations with the standardized birth rate. Sizable differences, including reversals of sign, occur in a number of instances, most notably in the correlations involving the measure of adult education. The lower panel presents the correlations involving the worker inputs. Comparison of the first two columns suggests fairly close agreement, despite the slight ambiguity in defining worker-inputs. (The third column of the lower panel will be discussed below.)

Table 8.4 shows the partial regression coefficients along with the means and standard deviations of the variables employed in the model. The regression model is shown for three different dependent fertility measures in 1968. Reflecting the differences in correlations and distributional parameters, notable differences in the regression coefficients occur in each of the three cases, as can be seen by comparing the columns labeled "Schultz" and "Schultz Replication". The pattern of difference is difficult to summarize: in every case, the magnitude of the child death and child schooling coefficients decrease in the replication; the effect of the male labor force becomes more positive in each instance in the replication, exhibiting reversal of sign in two cases; the magnitude of the VHEN worker-inputs is lowered in all three regressions; and the relative difference between the two types of workers changes greatly in two out of three cases. The t-statistics shown in Table 8.4 suggest that few of the differences would be statistically significant, though it might be noted that the labor force variable reaches significance at the 95 per cent level in only one of the two regressions (the original *or* the replication) for

Table 8.4. *Partial regression coefficients for Schultz linear regression model for crude birth rate, age-specific fertility 30-34 and 35-39, for 1968, compared with independent replication and effect of changing input measures (t-statistic given in parentheses)*

	Dependent variables								
	Crude birth rate, 1968			*A.S.F. 30-34, 1968*			*A.S.F. 35-39, 1968*		
Independent variables	*Schultz*	*Schultz replication*	*Hermalin input measures*	*Schultz*	*Schultz replication*	*Hermalin input measures*	*Schultz*	*Schultz replication*	*Hermalin input measures*
D_{1965}	3.95	3.514	3.491	3.98	3.734	3.758	12.1	11.103	11.031
	(13.4)	(15.135)	(14.976)	(8.12)	(8.419)	(8.425)	(12.5)	(12.567)	(12.437)
S_{1965}	−0.407	−0.336	−0.336	−0.676	−0.488	−0.485	−1.19	−0.843	−0.830
	(6.36)	(5.523)	(5.529)	(6.37)	(4.201)	(4.173)	(5.67)	(3.642)	(3.594)
E_{1965}	−0.030	−0.042	−0.045	0.038	0.044	0.049	0.857	0.754	0.755
	(0.30)	(0.555)	(0.586)	(0.23)	(0.301)	(0.334)	(2.62)	(2.600)	(2.609)
A_{1965}	0.032	0.085	0.086	−0.127	0.050	0.050	−0.280	0.267	0.270
	(0.88)	(2.000)	(2.014)	(2.10)	(0.614)	(0.615)	(2.35)	(1.643)	(1.665)
V_{1967} [a]	−0.011	−0.003	−0.003	−0.013	−0.006	−0.001	−0.067	−0.062	−0.039
	(0.89)	(0.284)	(0.485)	(0.65)	(0.297)	(0.068)	(1.70)	(1.495)	(1.548)
P_{1967} [a]	−0.001	−0.001	−0.001	−0.008	−0.006	−0.003	−0.029	−0.027	−0.017
	(0.33)	(0.565)	(0.857)	(1.83)	(0.004)	(1.222)	(3.46)	(3.046)	(3.170)
R^2	0.558	0.615	0.615	0.365	0.356	0.355	0.532	0.515	0.516
Mean	1.125	1.104	1.104	1.104	1.104	1.104	1.251	1.253	1.253
St. Dev.	0.228	0.213	0.213	0.315	0.315	0.315	0.724	0.723	0.723

Table 8.4 (supplement)

	Mean			Standard deviation		
	Schultz	*Schultz replication*	*Hermalin input measures*	*Schultz*	*Schultz replication*	*Hermalin input measures*
D_{1965}	1.061	1.063	1.063	0.033	0.036	0.036
S_{1965}	0.427	0.493	0.493	0.161	0.165	0.165
E_{1965}	0.424	0.528	0.528	0.088	0.115	0.115
A_{1965}	0.558	0.606	0.606	0.270	0.253	0.253
V_{1967} (*a*)	0.597	0.589	0.995	0.710 (*b*)	0.689	1.154
P_{1967} (*a*)	3.86	3.889	6.521	3.38 (*b*)	3.372	5.562

(*a*) For "Schultz Replication"—defined as cumulative worker-months of effort through 1967 per 1,000 women 15-49 as of June 1965. For "Hermalin Input Measures"—defined as cumulative worker-months of effort through 1967 per 1,000 married women 20-44.

(*b*) Obtained in private communication from T. Paul Schultz. Derivable from Table H-3 below.

Note: Dependent variables are normalized and the crude birth rate is also standardized. See Table 8.2 for definition of variables.

Sources: For "Schultz Regressions" see Schultz 1971a, Tables A-1, A-5, A-6 and Tables H-1, H-2, H-3. "Schultz Replication" and "Hermalin Input Measures" tabulations performed at Population Studies Center, University of Michigan.

each of the three dependent variables used. Whether changes of similar magnitude would occur for replications of the other dependent variables and for the other time periods reported by Schultz is not known.(9) The tables shown do indicate, however, that differentials in processing of the same basic data can be a source of non-trivial differences in the results of an analysis and that efforts to reduce this source of fluctuation are highly desirable. The results also suggest that the practice indicated in Table 8.2 for certain variables, of interpolating to obtain data for intermediate years, engenders the risk that undetected errors in the original observations will be carried over to the estimates. Where, as in this case, much of the data obtained by interpolation is available directly, it seems preferable to avoid estimation.

We turn next to the effect of varying the operational definition of a concept. It was noted earlier in this section that in their analysis of Taiwan, Schultz's measure of worker-input is per 1,000 *women* 15-49, while Hermalin's is per 10,000 *married women* 20-44. The

(9) It is possible that for 1965 and 1966 where Schultz (1971a) finds stronger effects from the worker-inputs, more substantial differences from the alternate data-processing might appear. It should also be noted that in this work Schultz finds that a non-linear model, incorporating second and third order terms of the worker-input measures, provides a better fit than the linear, and a test against this model would also be desirable.

differences resulting from these operational definitions are also displayed in Tables 8.2 and 8.3 (both measures are now expressed per 1,000). The bottom panel of Table 8.3, comparing columns 2 and 3, shows generally slight changes in the correlations of these two input measures with the other variables in the model. In Table 8.4, however, the effect of the definitional change on the regression results is more pronounced. Comparing for each dependent variable the "Schultz Replication" column with the "Hermalin Input Measures", where the only difference is the input variable, reveals the following:

(i) the regression coefficients for the non-input variables are little affected in every case;
(ii) the regression coefficients for the input variables are almost equal when the crude birth rate is the dependent variable, but the coefficients are sharply lower in magnitude when married women are used as the base in the other two cases.

From a statistical standpoint, the differences are more apparent than real. It will be noted from the supplement to Table 8.4, that the means and standard deviations of the Hermalin input measures are about two-thirds larger than the Schultz input measures, so that the coefficients of variation are about equal. Under these circumstances one would expect the corresponding regression coefficients to be two-thirds smaller using the Hermalin input measure and this is roughly the case when the dependent variable is age-specific fertility at ages 35-39. The same proportionate change does not occur for the other two dependent variables, though the effects here are statistically insignificant in both regressions. Thus, in this case, superficial examination of the effect of changing the input variables may lead to different conclusions about the returns from field-worker efforts which may not be warranted. In other cases, of course, the change may be more meaningful.

As before, there is no implication that one input measure is more correct than another. Arguments can be mounted in defense of either one, and a theoretical model that was sufficiently detailed might point quite unambiguously to the preferred operational definition. The point to be gained is the possible sensitivity of analytic results to this factor, as well as the other factors under discussion.

The third effect we wish to demonstrate is that arising from the inclusion or exclusion of a variable in a regression model, discussed in Section IV. In a preliminary version of his regression model, Schultz included a measure of cumulative doctor-input in addition to the variables shown in Table 8.4 The effect of this variable is

shown in Table 8.5, where the partial regression coefficients, with and without the doctor input variable, are compared for the model using the crude birth rate variable in 1968 as dependent variable:

Table 8.5. *Partial regression coefficients with and without doctor-input variable (Dependent variable = Crude birth rate, 1968)*

Independent variables	*Without Dr. input*	*With Dr. input*
D_{1965}	3.514 (*b*)	3.413 (*b*)
S_{1965}	−0.336 (*b*)	−0.309 (*b*)
E_{1965}	−0.042	−0.035
A_{1965}	0.085 (*b*)	0.059
V_{1967}	−0.003	−0.004
P_{1967}	−0.001	−0.008
Dr. Input (*a*)	—	−0.004

(*a*) Cumulative doctor-months of input through 1967 per 1,000 women, 15-49, as of June 1965.
(*b*) Regression coefficient is two or more times its standard error.
Source: Column 1 - from Table 4, column 2.
Column 2 - run at University of Michigan Population Studies Center.

The effect of including the doctor-input variable is to lower the magnitude of the coefficients of all the non-input variables and to raise the coefficients of the two worker-input measures, reversing their relative importance. Schultz (1971a; 26) argues convincingly that doctor inputs should be treated as an endogenous variable within a system of equations, since the number under contract to the program will be affected by the socioeconomic characteristics of an area, by the level of field-worker activity, and may be correlated with the disturbance term. The point of the illustration is not to dispute the specification used, but to show how the inclusion or exclusion of a measure, correlated with one or more of the independent variables, can affect the results.

We have attempted to provide numerical examples of some of the general points concerning the process of conceptualization and operationalization. Obviously, no general conclusions about the relative importance of each source of variation can be drawn from such select illustrations. Under differing circumstances, any of these sources can have either trivial or substantial effects. Of the three, the problem of variation arising from differential processing of the same source data is probably insufficiently appreciated, and merits

further attention. None of the problems discussed is unique to areal multivariate analysis, but may be more severe in this mode insofar as they are a function of small sample size. In the next section, we turn to two problems which are more particularly associated with the use of aggregate data.

VI. CHOICE OF AREAL UNITS AND SPATIAL AUTOCORRELATION

In this section we briefly treat two troublesome problems in areal analysis for which there are no clear-cut solutions. The first problem deals with the question of whether an analysis conducted at one level of aggregation will coincide with that at another level. Thus, one may use as units of analysis a smaller number of primary subdivisions or a larger number of secondary or tertiary subdivisions. In Taiwan, for example, published data may be given for 361 small administrative areas (chu, chen and hsiang) or for 21 areas representing the five largest cities in total plus 16 counties (hsien). The simple correlation of two variables based on the smaller areas will generally differ from one based on the larger areas since the former is a function not only of the latter correlation (weighted by the number of small areas per large area) but of the average correlation of small areas within the larger ones, and the correlation ratios of both variables (Duncan *et al.*, 1961 ; 110). A solution proposed by A. H. Robinson and discussed by Thomas and Anderson (1965), is to use weighted correlations or regressions where land areas are used as the weights. This approach, however, seems limited to those cases where the variables themselves are items like density which are based on areal measurements. For the types of variables being discussed here, the appropriate weighting variable will not always be apparent. A more serious limitation of Robinson's solution, pointed out by Thomas and Anderson, is that it produces equal parameter estimates across different levels of aggregation only under very restrictive conditions, namely, that the two variables have "exactly the same total distribution" under the two different modes of aggregation. Since this is rarely likely to occur, Thomas and Anderson suggest that one take "an inferential approach to the problem", and study whether the regression coefficients calculated at various levels of aggregation are significantly different. If they are not, the coefficients can be assumed to come from the same universe and that level of aggregation is not a factor in studying the particular relationships. The advice, in short,

is to examine the relationships of interest at different levels of aggregation and to note how they differ.([10])

Though there appears to be no unique solution to the problem of choosing the areal unit, the choice in practice may not be as troublesome as it appears in abstract. Unless the analyst contemplates creating aggregates out of small areal data, the published data will usually be available for only two or three levels. Given that the study of family planning effects will involve the multivariate analysis of a number of variables, one will usually want to avoid the largest level of aggregation—such as the states of India or the 21 hsiens and cities of Taiwan—since these provide too few cases for reliable estimation of parameters. Occasionally the structure of the family planning program may suggest the most "natural" level. For example, if decisions on inputs of program personnel, budgets, and the like, are made on the basis of factors at a certain areal level, this may be the most appropriate level for analysis.

In some cases problems may arise from too small a level of aggregation. A large number of areas with small populations may exhibit a great deal of random fluctuation from year to year in birth rates, death rates and other variables of interest. In such a case one may wish to aggregate over time (e.g., combine data for two or more years for each small area) or move to a higher level of areal aggregation, if this still provides a sufficient number of cases.

After choosing the type of areal unit to be used for analysis, a further decision involves the number of areas to be included. In some countries there may be some areas that are sufficiently distinct to provide grounds for their exclusion. In Taiwan, for example, of the 361 local administrative areas, 30 rural townships are designated as aboriginal because the non-Chinese inhabitants there differ in a number of ways from the rest of the population. They have substantially higher mortality and fertility, along with high educational attainment, and there was no input of field worker personnel into these areas before 1969. Since the total number of cases in areal analysis is rather small, the inclusion or exclusion of a number of distinct units can have a noticeable effect on a regression analysis. Thus, Schultz (1971a; 30) found that his measure of total personnel impact on birth rates declined when these areas were omitted.

Lastly, it may be noted that the designated areal units may be

([10]) Though weighted regressions will not necessarily ensure similar results at two different levels of aggregation, weighting is commonly employed in areal analysis (see Nerlove and Schultz, 1970 and Collver *et al.*, 1967). Whether weighted or unweighted regressions should be employed must be decided on conceptual as well as statistical grounds.

classified in various ways such as urban-rural; location (east or west of a natural barrier, etc.) or fertility prior to the family planning program. Results of a multivariate analysis may then be obtained for each subset in order to examine the uniformity of the patterns; alternatively, the classification itself can be incorporated into a regression analysis via dummy variables, to study its effect.

The second problem alluded to at the outset arises from the tendency of adjacent areas to resemble one another more closely than those further apart. As a result, one of the assumptions of regression analysis, noted in Section IV, that the disturbance terms be uncorrelated with one another, may not hold when areas are the unit of analysis. The problem is analogous to that of autocorrelation in time series and is sometimes referred to as spatial autocorrelation. A review of various approaches that may be taken is given by Berry (1971) and Duncan *et al.* (1961; 128-160). Measures designed to determine whether spatial autocorrelation exists in the original data or in the regression residuals may prove of some assistance in gauging the severity of the problem, but Berry finds that the measures proposed "are invariant under certain topological transformations... and dependent upon the criteria for assessing whether or not counties are contiguous". In general, solutions that are useful in handling time series problems are less applicable to spatial autocorrelation because the assumptions of stationarity are more tenable with temporal than with spatial phenomena. Under this circumstance, the results of a regression analysis based on areal data must be interpreted with some caution.

The bias from spatial autocorrelation may be viewed as a form of specification error. If adjacent areas are more alike than those further apart in terms of the residuals from a given model, this indicates that all the factors accounting for areal difference have not been incorporated into the model. The addition of these missing variables can be expected to alter the observed regression coefficients, unless the omitted variables are uncorrelated with those included (Duncan *et al.*, 1961; 147).

VII. MULTIVARIATE TECHNIQUES OF ANALYSIS

The variables ordinarily used in an areal analysis will be measured as ratio or interval scales, so that multiple regression and factor analysis techniques are potentially applicable. This section discusses the use of these techniques in estimating the parameters of a model.

It was noted in Section IV that a conceptual model may, more

or less explicitly, account for more than one variable, and the exposition of these interrelationships gives rise to a series of equations. In such a model, one can distinguish the exogenous (or predetermined) variables which the theory does not seek to explain, from the endogenous (or jointly dependent) variables whose values are accounted for. The areal analysis by Hermalin described in Section III can be represented by the following set of equations and the diagram in Figure 8.1.([11])

$$Y_1 = P_{12} Y_2 + P_{13} Y_3 \tag{2}$$

$$Y_2 = P_{25} Y_5 + P_{2a} X_a + P_{2c} X_c + P_{2u} R_u \tag{3}$$

$$Y_3 = P_{34} Y_4 + P_{35} Y_5 + P_{3b} X_b + P_{3c} X_c + P_{3d} X_d + P_{3w} R_w \tag{4}$$

$$Y_4 = P_{45} Y_5 + P_{4a} X_a + P_{4c} X_c + P_{4v} R_v \tag{5}$$

$$Y_5 = P_{5a} X_a + P_{5c} X_c + P_{5z} R_z \tag{6}$$

where,

Y_1 = logarithm of age-specific fertility in year t.([12])
Y_2 = logarithm of age-specific proportion married in year t.([12])
Y_3 = logarithm of age-specific marital fertility in year t.([12])
Y_4 = age-specific cumulative acceptance rate through first quarter of year t.
Y_5 = age at marriage (measured as average proportion married at 20-24 in 1961 and 1963; a measures which correlates—.90 with singulate mean age at marriage).
X_a = level of adult education in 1963; per cent of females 12 and over, not in school, who are primary school graduates or higher.
X_b = level of child schooling in 1965; per cent of females 12-19 currently attending school.
X_c = per cent of male labor force (12 and over) in agriculture and fishing, in 1963.
X_d = child mortality; weighted average of death rates 0-4, in 1961 and 1963.

The equations and the diagram utilize the conventions of path analysis, a multivariate technique useful in explicating linear causal models. It bears close affinities to multiple regression analysis but helps make explicit the underlying assumptions and interconnections. More detailed expositions of this techniques can be found in Duncan (1966) and Land (1969). This presentation draws heavily on Duncan.

([11]) The presentation here is more detailed in some respects than in Hermalin (1971) where the focus was on only the first three equations.

([12]) The data to be presented here are for women 30-34 in 1968.

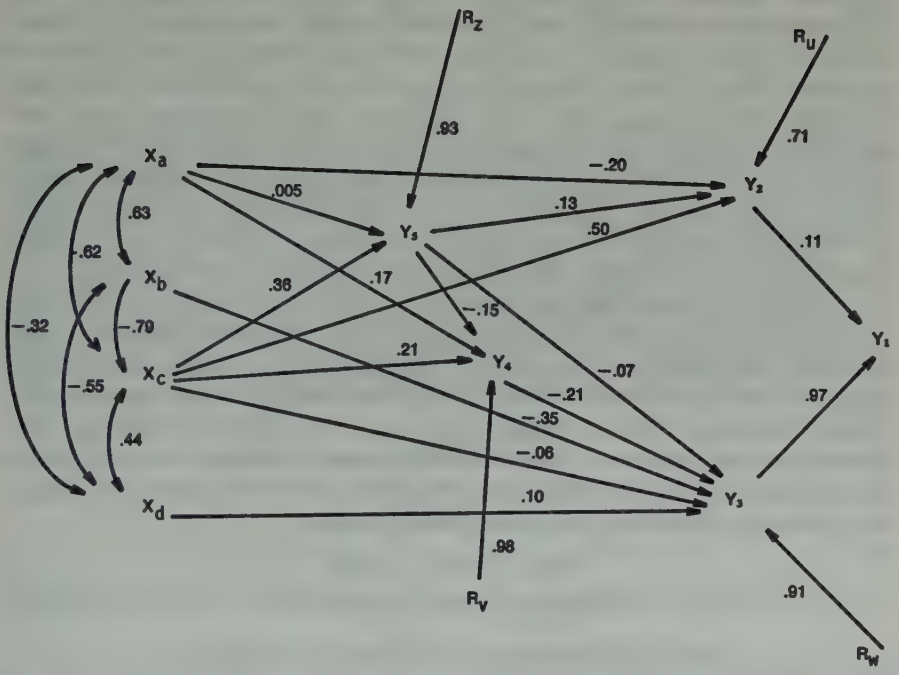


Fig. 8.1. *Path diagram of factors affecting age-specific fertility through the components proportion married and marital fertility*

Note: The path coefficients from the residuals are calculated as $\sqrt{1-R^2}$, where R^2 is the coefficient of determination for the respective dependent variables (i.e., the proportion of variance accounted for by the explanatory variables with paths to the dependent variable). (See Table 8.6.)

In the equations, each variable is in standard form; the Y_i variables are considered endogenous, the X_i exogenous, and the R_i the residual effects of unmeasured variables, which are uncorrelated with the other determining variables in each equation. Thus, each dependent variable is treated as completely determined by some set of other variables in the model, including where necessary the unmeasured residuals. The P_{ij} are path coefficients representing the direct effect on variable i from variable j. The variables in Figure 8.1 are arranged in a temporal and presumed causal sequence.([13]) Each variable occurs earlier in time than those appearing to the right of it and later than those to the left of it, and can be affected by all the variables that precede it. The direct or net influences from one variable to another are represented by the one-way straight arrows and the sign and magnitude associated with these are the path coefficients. The curved, bi-directional arrows represent the unanalyzed correlations between the exogenous variables in the model, and the values shown with these are the simple correlation coefficients.

It will be noted that equation (2) above differs from the others in that no residual term is included for complete determination. This equation represents the algebraic decomposition of age-specific fertility into two components—the proportion married and age-specific marital fertility. Where illegitimacy is not a significant factor and where the calculation of marital fertility rates is based on total births (legitimate and illegitimate), both of which conditions hold in Taiwan, the following relations are algebraically exact:

$$ASF = PM \times ASMF \tag{7}$$

$$\ln ASF = \ln PM + \ln ASMF \tag{8}$$

Table 8.6. *Correlation matrix of variables in figure 8.1*

	Y_2	Y_3	Y_4	Y_5	X_a	X_b	X_c	X_d	*Mean*	*Standard deviation*
Y_1	0.260	0.994	−0.230	0.160	−0.200	−0.406	0.292	0.289	2.22	0.11
Y_2		0.153	0.078	0.352	−0.537	−0.626	0.670	0.424	−0.02	0.01
Y_3			−0.245	0.122	−0.142	−0.343	0.220	0.247	2.24	0.11
Y_4				−0.116	0.072	0.077	0.050	−0.084	342.00	90.29
Y_5					−0.218	−0.409	0.357	0.488	61.73	9.44
X_a						0.631	−0.618	−0.318	38.83	13.25
X_b							−0.792	−0.548	37.27	14.71
X_c								0.437	59.03	25.28
X_d									12.17	3.79

([13]) This is not strictly true of variables Y_1, Y_2 and Y_3 [equation (2)], which are contemporaneous and connected by an algebraic identity.

Path coefficients for specified combinations of variables

endent iables	Independent variables								R^2 Coefficient of determination
	X_a	X_b	X_c	X_d	Y_5	Y_4	Y_3	Y_2	
	0.005	—	0.36	—	—	—	—	—	0.13
	0.17	—	0.21	—	−0.15	—	—	—	0.04
	—	−0.35	−0.06	0.10	−0.07	−0.21	—	—	0.17
	−0.20	—	0.50	—	0.13	—	—	—	0.49
	—	—	—	—	—	—	0.97	0.11	

The model thus examines the relative contribution of each component of age-specific fertility and the relative importance of various factors in the determination of each component. It also permits one to study how the effects of these factors are transmitted to age-specific fertility through its components, which will be taken up below. In the case of an algebraic identity, the path coefficients in equation (2) can be obtained directly as the ratio of the standard deviation of each component to the standard deviation of the dependent variable. If $\tilde{Y}_1$, $\tilde{Y}_2$, and $\tilde{Y}_3$ represent the unstandardized variables, then

$$\tilde{Y}_1 = \tilde{Y}_2 + \tilde{Y}_3 \tag{9}$$

$$\bar{Y}_1 = \bar{Y}_2 + \bar{Y}_3 \tag{10}$$

$$\bar{Y}_1 - \tilde{Y}_1 = (\bar{Y}_2 - \tilde{Y}_2) + (\bar{Y}_3 - \tilde{Y}_3) \tag{11}$$

Multiplying and dividing each term by its standard deviation gives

$$\frac{\bar{Y}_1 - \tilde{Y}_1}{\sigma_1}\sigma_1 = \frac{\bar{Y}_2 - \tilde{Y}_2}{\sigma_2}\sigma_2 + \frac{\bar{Y}_3 - \tilde{Y}_3}{\sigma_3}\sigma_3 \tag{12}$$

and dividing each term by σ_1, we have

$$\frac{\bar{Y}_1 - \tilde{Y}_1}{\sigma_1}\frac{\sigma_1}{\sigma_1} = \frac{\bar{Y}_2 - \tilde{Y}_2}{\sigma_2}\frac{\sigma_2}{\sigma_1} + \frac{\bar{Y}_3 - \tilde{Y}_3}{\sigma_3}\frac{\sigma_3}{\sigma_1} \tag{13}$$

which is equivalent to

$$Y_1 = \frac{\sigma_2}{\sigma_1}Y_2 + \frac{\sigma_3}{\sigma_1}Y_3 \tag{14}$$

where Y_1, Y_2, and Y_3 are in standard form. Comparison with equation (2) shows that:

$$P_{12} = \frac{\sigma_2}{\sigma_1} \quad \text{and} \quad P_{13} = \frac{\sigma_3}{\sigma_1} \tag{15}$$

The model in Figure 8.1 does not posit any reciprocal causation or feedback between variables. This can also be seen from the basic equations (2) through (6) in that there is no instance where both P_{ij} and P_{ji} appear. Such a system is termed recursive. Under the assumption that the residual effects are uncorrelated with each other and with the exogenous variables in the model, the system of equations can be solved equation by equation by ordinary least squares regression.([14]) In this case, the path coefficients will be equal to the partial regression coefficients in standard form, or Beta coefficients, of the regression results. In this model, for example, focusing on equation (6), $P_{5a} = \beta_{5a.c}$ and $P_{5c} = \beta_{5c.a}$, obtained from the regression of Y_5 on X_a and X_c.

Alternatively, one may solve for the unknown path coefficients in terms of the observed correlations. Since the variables are in standard form,

$$r_{5a} = \frac{\sum X_a Y_5}{N} \tag{16}$$

Substituting for Y_5 from equation (6) gives

$$r_{5a} = \frac{1}{N}\sum X_a (P_{5a} X_a + P_{5c} X_c + P_{5z} R_z) \text{ or} \tag{17}$$

$$r_{5a} = P_{5a} + P_{5c} r_{ac} \tag{18}$$

since

$$\frac{\sum X_a X_a}{N} = 1 \quad \text{and} \quad r_{az} = 0$$

Similarly,

$$r_{5c} = \frac{\sum X_c Y_5}{N} \tag{19}$$

and the same process yields

$$r_{5c} = P_{5a} r_{ca} + P_{5c} \tag{20}$$

Equations (18) and (20) can now be solved simultaneously for the path coefficients P_{5a} and P_{5c} by using the three observed correlations contained in the equations.

This expansion of the correlations into path coefficients illustrates the "basic theorem of path analysis" which may be written generally as

$$r_{ij} = \sum_q P_{iq} r_{jq} \tag{21}$$

([14]) For the model shown in equations (2) through (6), since it is not fully recursive, it is not necessary to posit zero correlation among all the residuals. For example, R_u and R_w may be correlated, since Y_2 does not appear in equation (4) and Y_3 does not appear in equation (3).

where i and j are two variables and q runs over all the variables which have paths leading directly to variable i (Duncan, 1966, 5-6).

The model in Figure 8.1 does not posit effects from all antecedent variables to all subsequent ones. It is assumed, for example, that only X_a and X_c affect Y_5 directly, but that there is no direct effect from X_b or X_d to Y_5. The adequacy of the assumptions can be tested by comparing the correlations r_{5b} and r_{5d} implied by the model with those observed from the data.

The implied correlations can be calculated from the basic theorem as follows:

$$r_{5b} = P_{5a} r_{ab} + P_{5c} r_{bc} \tag{22}$$

$$r_{5d} = P_{5a} r_{ad} + P_{5c} r_{cd} \tag{23}$$

Using the values given in Figure 8.1 or from Table 8.6, the implied correlations are $r_{5b} = 0.282$ and $r_{5d} = 0.156$. The observed correlations are (0.409) and (0.488) respectively, and the sizable discrepancy suggests that the model would be improved by including direct effects to Y_5 from X_b and X_d. Unless the theoretical model specifies the precise interconnections, an alternative approach for determining which paths can be omitted is to first calculate the effects from all the antecedent variables. Those that are trivial in magnitude and/or statistically insignificant can be dropped from the model and the path coefficients recalculated for the direct paths retained.

These results illustrate two possible uses of path analysis in relation to theory development. As Heise notes (1969; 62-63) a *full-scale* recursive system, where all possible causal relationships are included, will always reproduce the observed correlations. Given two conceptual models which agree on the terms of the basic temporal or causal ordering, but which differ in the paths deleted, their relative ability to reproduce the observed correlations can be tested and one, or both, of them rejected if there is poor agreement. A related procedure which Heise calls "Theory Trimming" starts with the full-scale recursive system and then deletes those paths which are negligible, in the interest of obtaining the most parsimonious model which will reproduce the correlations. Useful as these procedures may be, it must be cautioned that they do not provide a test of the basic causal ordering of the variables.

The way in which the effects of the variables in Figure 8.1 are transmitted to age-specific fertility via its components can also be explicated by use of the basic theorem. We illustrate this for the variable Y_5—the age at marriage proxy—as follows:

$$r_{15} = P_{12} r_{25} + P_{13} r_{35} \tag{24}$$

Using the values from Table 8.6, this produces:

$$0.160 = 0.11\,(0.352) + 0.97\,(0.122)$$

$$0.160 \cong 0.039 + 0.118$$

Thus, earlier age at marriage (leading to higher values of Y_5) increases age-specific fertility primarily through its effect on marital fertility rather than through proportion married at 30-34. If desired, the correlations on the right-hand side of equation (24) can be further decomposed by additional application of the basic theorem.

Path analysis is also useful in explicating the indirect effects inherent in the model. The correlation between two variables can be decomposed into a direct effect, indirect effects, and the joint effects shared with other variables in the system (Duncan, 1971; 137-138). With reference to Figure 8.1, the correlation between Y_2 and X_c may be decomposed as follows:

$$r_{2c} = P_{2c} + P_{25}\,r_{5c} + P_{2a}\,r_{ac} \tag{25}$$

and upon further expansion, we have

$$r_{2c} = P_{2c} + P_{25}\,(P_{5c} + P_{5a}\,r_{ac}) + P_{2a}\,r_{ac} \tag{26}$$

or

$$r_{2c} = P_{2c} + P_{25}\,P_{5c} + r_{ac}\,(P_{2a} + P_{25}\,P_{5a}) \tag{27}$$

The first term on the right side represents the direct effect of X_c on Y_2, the second term, the indirect effect of X_c on Y_2 via Y_5 and the last term the effect on Y_2 which X_c shares with X_a.

Substituting the values from Table 8.5, we have:

$$0.67 = 0.50 + 0.05 + 0.12$$

In summary, then, path analysis is a useful multivariate technique for explicating the interrelationships of a conceptual model; under certain circumstances it can reveal the relative utility of two conceptual models in accounting for a set of observed correlations and lead to a more parsimonious representation of these correlations.

In situations where the assumption of uncorrelated disturbances across equations does not hold, the application of ordinary least squares regression will yield biased and inconsistent estimates of the equations' parameters. Consider equations (5) and (6) of the model used at the beginning of this section:

$$Y_4 = P_{45}\,Y_5 + P_{4a}\,X_a + P_{4c}\,X_c + P_{4v}\,R_v \tag{5}$$

$$Y_5 = P_{5a}\,X_a + P_{5c}\,X_c + P_{5z}\,R_z \tag{6}$$

If the unmeasured factors affecting Y_4, the cumulative acceptance rate, are correlated with the unmeasured influences on Y_5, age at marriage, this means that R_v and R_z will be correlated. Since R_z determines in part Y_5, the values of Y_5 in equation (5) will be correlated with R_v, violating the condition of ordinary least squares that the explanatory variables be independent of the disturbance term. (This can be seen algebraically by multiplying equation (6) by R_v, summing and dividing by N. This shows that r_{5v} is a function of r_{vz} and insofar as this latter correlation is not zero, the conditions for using ordinary least squares in equation (5) will not be met.)

The assumption of correlated disturbance terms across equations in the case of a system of the type under discussion may arise from knowledge of the nature of the unmeasured factors and their presumed influences. In addition, one may calculate the correlations among estimated residuals and examine their magnitude. Schultz (1971b; 33-37) carries out a test of such correlations, based on Zellner's (1962) estimation procedures for "seemingly unrelated regressions".

If the model posits reciprocal influence between two or more endogenous variables, then the conditions for using ordinary least squares are not met, because of the simultaneity involved. Assume in the example above that in equation (6), Y_5 is affected by Y_4 as well as the exogenous variables shown and that equation (5) remains as given. Thus, Y_4 and Y_5 now affect each other. If equation (5) is multiplied by R_z and equation (6) by R_v, and sums taken, it will be seen that, unless the reciprocal influences are zero, Y_4 will be correlated with R_z and Y_5 with R_v, violating the condition of using ordinary least squares to estimate equations (5) and (6).

Econometricians have developed a number of estimation techniques which can be used in lieu of ordinary least squares (see for example J. Johnston, 1963, Chapter 9). Some of these techniques can be applied within the framework of path analysis as shown by Duncan *et al.* (1968b) but this approach will not be pursued here. Of the techniques available other than ordinary least squares regression, we discuss only the frequently used method of two-stage least squares. The main idea is to "purge" the endogenous variables, used as explanatory variables in the system, of their correlation with the disturbance terms. This is accomplished by using estimates of the endogenous variables based on a first-stage regression on all or some of the exogenous variables in the model (Johnston, 1963; 236-237). Using the two equations above as an example, Y_5 is first regressed on X_a and X_c (and if desired X_b and X_d as well). The estimate of Y_5 thus obtained, $\hat{Y}_5$, is then used in place of Y_5 in

equation (5). Since $\hat{Y}_5$ is now a linear combination of exogenous variables, and exogenous variables are assumed independent of the disturbance term, it follows that $\hat{Y}_5$ will be independent of the disturbance term and equation (5) can be estimated by ordinary least squares. In order to apply this technique to any equation, the condition of identifiability must be met. This requires that the number of exogenous variables excluded from any equation be at least equal to the number of right-hand-side endogenous variables included. In equation (5), for example, two exogenous variables, X_b and X_c are excluded and there is only one right-hand-side endogenous variable, Y_5, so that two-stage least squares is applicable. Of the 5 equations, (2) to (6), equation (4) would not meet the identifiability criterion. Though two-stage least squares, and related techniques, provide estimates which are asymptotically unbiased, i.e., consistent, their small-sample properties are biased (Johnston, 1963; 234-237, 275). Monte Carlo studies of the behavior of various estimating techniques are reviewed by Johnston (1963; chapter 10).

In Section III and Table 8.1, Schultz's (1971b) simultaneous equation model for Taiwan was reviewed. Schultz presented estimates obtained by two-stage least squares. For the equation determining age-specific fertility rates he found in almost all comparisons that the effect of cumulative IUD acceptance rates is much lower under the two-stage procedure. The two multivariate techniques can be expected to yield different estimates, but the degree of difference will be a function of the conceptual model employed and the operational definitions.

As an example of this, Table 8.7 presents the regression coefficients of IUD acceptance rates from each technique in the equation determining age-specific fertility rates from Schultz's model, and those obtained from an alternate simultaneous equation model, developed by this writer. In Hermalin's model, as in Schultz's, the use of two-stage least squares tends to lower the apparent effect of acceptances, but the change is more moderate than in Schultz's model, and the effect remains significant in two of the four years. Comparisons of ordinary least squares and two-stage least squares carried out for other conceptual models of the determinants of fertility, not presented here, lead to the same conclusion: that the differences between estimation techniques will be sensitive to model specification, including differences in the instrumental variables used to estimate the endogenous factors.

Heretofore we have presented results from multiple regression and path analysis in which both standardized regression coefficients (Beta coefficients) and unstandardized coefficients have been utilized.

Table 8.7. *Comparison of partial regression coefficients from ordinary least squares and two-stage least squares solutions for two models concerning determinants of fertility in Taiwan*

Partial regression coefficients of IUD acceptance rates in fertility equation, ordinary least squares vs. two-stage least squares

Model	OLS		TSLS	
	b	*t*	*b*	*t*
Schultz II				
ASF 15-19: 1966	−0.027	3.69	−0.003	0.13
1968	−0.003	0.74	0.003	0.14
ASF 20-24: 1966	−0.008	3.64	−0.004	0.62
1968	−0.000	0.02	0.004	0.72
ASF 25-29: 1966	−0.000	0.33	0.006	0.85
1968	−0.000	0.55	−0.011	1.24
ASF 30-34: 1966	−0.004	1.23	0.002	0.11
1968	−0.009	4.01	−0.002	0.14
ASF 35-39: 1966	−0.021	3.26	0.080	1.30
1968	−0.019	4.15	0.049	0.95
ASF 40-44: 1966	−0.011	1.02	0.014	0.21
1968	−0.018	2.57	0.002	0.03
ASF 45-49: 1966	−0.029	1.32	0.026	0.27
1968	−0.019	1.45	−0.150	1.54
Hermalin II				
TFR: 1966	−3.168	5.22	−2.779	2.10
1967	−2.164	4.79	−3.258	3.49
1968	−1.726	4.72	−1.319	1.66
1969	−1.251	3.78	−0.739	1.12

Schultz II: ASFi: per cent married*; cumulative acceptance rates*; doctor input*; child mortality measure; child education; adult female education; male income; agriculture and fishing labor force; VHEN worker input; PPH worker input; relative education male/female, N = 361.
Hermalin II: TFR: per cent married*; cumulative acceptance rate*; child mortality, N = 331.
* Treated as endogenous in TSLS.
Note: The partial regression coefficients in Schultz II and Hermalin II are not directly comparable because of differences in the units of measurement for the independent and dependent variables. Also, the t-statistics from two stage least squares are not directly comparable to those from ordinary least squares.
Sources: Schultz II—Schultz, 1971b, Table 8a. See also Section 2 and Table 8.1 of this paper. Hermalin II—unpublished tabulations of the Population Studies Center, University of Michigan.

The unstandardized coefficients give the amount of change in the dependent variable associated with a one unit change in the independent variable, where each variable is measured in its original units. Coefficients in this form are particularly helpful in studying the change expected in acceptance rates or fertility with increases in various types of inputs and thus form a basis for decision-making

with respect to the family planning program. The standardized regression coefficients may be interpreted as the amount of change in standard deviation units in the dependent variable associated with a one standard deviation change in the independent variable, with the other variables held constant.

The relation of the two types of coefficients is simply expressed as:

$$P_{ij} = b_{ij} \frac{s_j}{s_i} \tag{28}$$

where
P_{ij} is the standardized coefficient,
b_{ij} is the unstandardized coefficient,
s_j and s_i are the standard deviations of the independent and dependent variables respectively (Schoenberg, 1972a; 3-4). The equation shows that if there are differences in the standard deviation, the standardized coefficients may differ across two samples even if the unstandardized coefficients agree. The dependence of the standardized coefficients on the standard deviations has led Schoenberg to suggest the following guidelines:

> "... the path of variable j to i can be compared with the path of variable k to i if the paths are computed from the same population, since they are contributing to the same variance in variable i. However, the same path can not be compared with any other path that does not impinge on the same dependent variable i, not even if that other variable is identically measured on another population. It is simply not legitimate to compare standardized measures across different samples since any difference in such measures may just as likely be attributable to differences in variances as to differences in effects... paths may be compared only with those other paths that impinge on the same dependent variable (within the same set of data) to which it contributes variation." (Schoenberg 1972a; 4.)

Blalock (1967; 675) suggests that standardized or path coefficients are "appropriate if one wishes to measure the *actual* amount of impact that each variable has on the other in a given population" (italics in original) while the unstandardized form is more appropriate for describing causal structure. In general, then, when one wants to go beyond describing a given time and place and compare effects across years, across subsamples in a country, or across countries, unstandardized regression coefficients are to be preferred.

Another method of multivariate analysis applicable to areal data is that of factor analysis, developed largely in the field of psychometrics. It seeks to reproduce a data matrix, or matrix of correlations, from a smaller set of unobserved or latent factors. The focus is on the mutual interdependence of a set of variables rather than on

the dependence of a given variable on a number of explanatory variables. The mathematics of factor analysis is given by Fruchter (1954), Harmon (1967) and Rummel (1970). The factor analysis solution also indicates the degree to which the variance of each variable is accounted for by each factor extracted from the correlation matrix. (The square root of the proportion of variance accounted for is called the factor loading.) As applied to an areal analysis of fertility, a measure of fertility along with a number of other variables for each area can be factor analyzed. For each factor the extent to which it acounts for the fertility variable can be noted, along with the other variables mainly accounted for by this factor. If these other variables hang together in terms of common content, the nature of the factor can be inferred. For example, in an analysis of the fertility of Honolulu using census tracts as the unit of analysis, Park (1972) found that the first common factor accounted for a high proportion of the fertility measure and a number of other variables, such as female labor force, male labor force, male unemployment, having to do with proportion of the population which is economically active. He found that the other common factors similarly identified clusters of variables suggesting other socioeconomic aspects of the tracts. Park also compared the implications of the factor analysis with that from a multiple regression analysis of fertility.

A somewhat different use of factor analysis was employed by Le Bras (1971). His aim was to study the fertility differences among the 89 départements of France. To this end he constructed a data matrix which showed the "fertility distance between any two departments." This "distance" was calculated as the square root of the sum of the square differences of the seven age-specific fertility rates, a measure which Le Bras felt captured both the structure and intensity of fertility differentials. A principal components factor analysis revealed that the variance of the data matrix was almost entirely captured by two orthogonal axes or factors. A plot of the loading of each departement on each factor indicated that the departements tend to cluster by region. In addition, Le Bras shows that the first factor was associated with the gross rate of reproduction while the second was reflective of the mean age of childbearing.

Factor analyses of the type described do not produce unique solutions. Mathematical criteria, such as orthogonality, the number of factors to be extracted and the pattern of zero and non-zero loadings, rather than theoretical criteria, are employed to attain a solution. It is to be noted that these analyses are not motivated by a prior conceptual model of the underlying dynamics, and these approaches may be termed exploratory factor analysis. [For a dis-

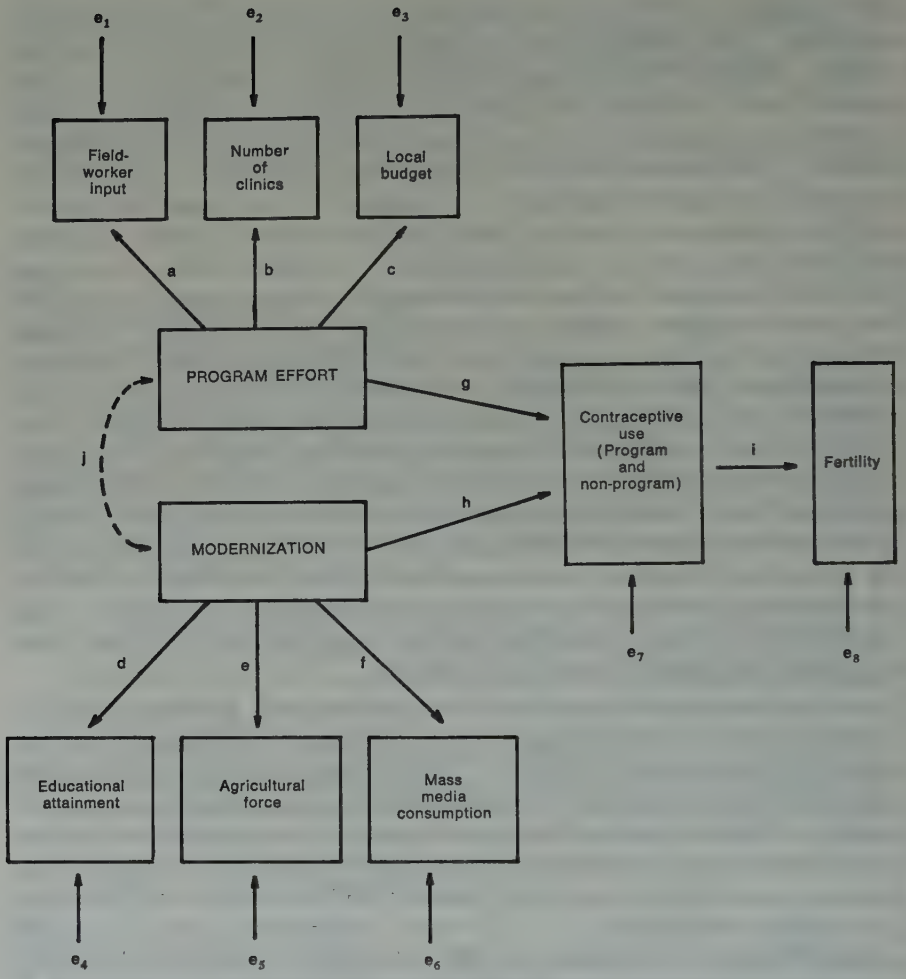


Fig. 8.2. *Hypothetical model for studying family planning program impact using unmeasured variables and multiple indicators.*

cussion of exploratory vs. confirmatory factor analysis, see Jöreskog and Lawley (1968).]

Factor analysis can also be used in conjunction with conceptual models. In Section V, the desirability of including unobservable concepts explicitly in the analysis via their indicators was pointed out. In this confirmatory use of factor analysis, the variables thought to be causes or effects of the unobserved factor are specified in advance, and goodness of fit tests can indicate the appropriateness of the model.

Figure 8.2 presents a hypothetical model using multiple indicators to study the relative impact of a family planning program where "program effort" and "modernization" are treated as unmeasured variables. No claim is made that all the parameters of such a model can be identified. (See Werts *et al.*, 1973; 1473-1474.) The model given can be substantively modified in numerous ways, and can also be modified to include assumptions about correlated measurement errors. As in the examples concerning recursive systems, techniques for handling unmeasured variables and multiple indicators cannot establish the correctness of a given model. Under certain circumstances they can indicate that a certain model is faulty and, through goodness of fit tests, indicate which of several models best accords with the data (Duncan, 1972; 40-42; Werts *et al.*, 1973).

The multiple indicator approach appears to be a promising development in the areal analysis of family planning program effects, though this writer is not aware of any such studies in which it has been used. It is supportive of more sophisticated model building and permits a greated concordance between theoretical constructs and empirical analysis. In conventional multiple regression, one runs the risk of specification error by excluding variables which have explanatory impact on the dependent variable and which are correlated with the independent variables. At the same time, the inclusion of a larger number of correlated variables may lead to the problem of redundancy. As Gordon (1968) has shown, when there are clusters of interrelated variables, the apparent effect of any one variable may be a function of the number of variables in any one set, in that the clusters tend to divide up their joint influence among the individual members. Lastly, the strategy of using only a limited number of variables from a cluster followed by Hermalin (1971; 19-20) may provide a less than adequate measure of the underlying construct. The use of unmeasured variables and multiple indicators largely avoids these problems (Schoenberg, 1972b). [For further references on multiple indicator models, see also Hauser and Goldberger (1971) and Schoenberg (1972b).]

VIII. INTERPRETATION AND USE OF RESULTS

The multivariate areal models illustrated in this paper are essentially cross-sectional in nature and dynamic inferences do not readily follow. That is, the models investigate the extent to which differentials in fertility across areas at a certain date are associated with other characteristics of the area, some of which may be measured at the

same time and others for preceding points in time. Despite the use of lagged variables, "a finding of influence from one variable to fertility does not mean that changes in this variable within an area are associated with intra-area changes in fertility over the time period specified by the model; nor does the absence of association mean that the background variable is without influence on intra-area fertility changes" (Hermalin, 1971; 23).

Hermalin (1971; 24) illustrates the first point as follows:

> "Assume that the areas at the start of the period show a strong positive association between child mortality and fertility; and that thereafter child mortality remains constant while fertility declines at an equal rate across areas. In this case, the model would show a high association between child mortality and later fertility but it would be incorrect to regard the *change* in fertility as due to a *change* in mortality over the period covered. The noted association existed at the start of the period and is likely the result of a long historical process; under the assumption given, it cannot be due to *change* in mortality over the life of the program."

As illustration of the second point, Hermalin provides an example where all the areas are assumed equal in child mortality at the start of the family planning program. This variable then would show no effect on fertility according to the model. However, the areas, over the period covered by the lag, could experience both a decline in mortality and fertility, leading to a different inference about the effect of mortality (Hermalin, 1971; 24).

Schultz (1973; S257), in discussing the same problem, summarizes it as follows:

> "The estimates from cross sections using discretely lagged levels will tend to understate long-run responses because a single annual observation approximates the appropriate weighted distribution of past observations with substantial error. On the other hand, a discrete 3-year lag overstates short-run responses and creates the erroneous impression that the estimated response would occur in three years, whereas in all likelihood it would take longer."

Additional aspects of the relation between spatial and temporal variation are given by Duncan *et al.*, (1961; 160-174).

Improved estimates of dynamic relationships are possible when one has cross-sections for several years. A method which utilizes pooled time series of cross-sections is briefly described in Nerlove and Schultz (1970; 37-42) along with its application to a study of Puerto Rican fertility. The approach treats the disturbance terms of the regression as containing two components, one which is specific to an area and relatively time invariant and a second which is both area and time independent.

Schultz (1973) has used this technique with pooled cross-sections in Taiwan from 1964 to 1969 and finds that in comparison with the single cross-section estimates, the pooled data indicate smaller effects from child mortality and the socio-economic variables and about double the effect from field-worker inputs. He also employs a first difference model between fertility levels in 1969 and those in 1964 as a way of removing the area-specific time-invariant effects which can bias the results obtained in single cross-sections. Certain limitations associated with correlations and regressions based on differences are noted in Duncan *et al.* (1961; 162-166) and Freedman and Takeshita (1969; 330-331). Where cross-sections for several years are available, the pooled cross-section approach appears to be a promising way of improving the temporal specification of a model.

Another limitation in the interpretation of areal analysis was noted at the outset of this paper, namely, that relationships found to hold among areal units cannot be assumed to hold among individuals. However, if the problem has been properly defined as one concerning aggregate behavior, this poses no restriction on the use of areal data.

Various uses of areal analysis have been illustrated in the course of this paper, in addition to the general objective of determining whether a family planning program has had a significant impact on fertility, while taking account of other factors. Policy decisions can be aided by studying the returns obtained from different types of inputs; the differential success of the program in areas with differing characteristics may suggest a reallocation of effort. Given the limitations associated with areal analysis stressed in this paper, it is wise to seek confirmatory evidence from other sources, but an on-going program must proceed on the basis of the best available imperfect knowledge, while the search for better information continues.

Confirmatory evidence for the inferences from an areal analysis can come from many sources. Alternate approaches to studying the impact of family planning programs reviewed in this volume provide an obvious check. Another, when appropriate, can come from testing data on individuals obtained from KAP or other surveys. Yet a third avenue is provided by seeing whether alternate statistical procedures, involving different assumptions about the nature and form of the data, lead to the same conclusions. Another type of safeguard arises from a close knowledge of both the data-producing operations in a country and their demographic, social and economic trends. These insights may reveal changes in statistical procedures, important but temporary phenomena, basic changes in law, etc. which may

affect fertility (or reported fertility) and confound the results of a short-range statistical analysis.

Results of an areal multivariate analysis are strengthened when the same implications are obtained from alternate models and/or estimation procedures. As an example, Schultz's (1973) finding from a first-difference model, that field workers in Taiwan had greater impact in those areas that had a higher than expected fertility (on the basis of positive residuals) is generally confirmatory of Hermalin's (1971) earlier finding, based on single cross-sections with a different specification, that the program was more successful in areas with higher pre-program fertility.

The viewpoint taken in this paper is that important questions concerning the impact of a family planning program on fertility are amenable to study through areal multivariate analysis. Despite this appeal, a number of limitations which this mode of analysis shares with other multivariate studies has been stressed. Results are likely to vary as a result of differences in conceptual models, operational definitions, processing of data, and techniques of estimation. In addition, problems connected with the choice of areal unit and spatial autocorrelation have been noted. The implicit approach taken has been to view multivariate areal analysis as an example of causal analysis and to present in rather simplified form aspects of this larger subject as it pertains to this particular case. Multivariate areal analysis will clearly benefit from better theory of fertility behavior, both on the individual and aggregate level, from greater standardization in operational definitions, reduction in measurement error, and more efficient estimating techniques. Several promising developments were pointed out. The techniques and concepts employed in causal and areal analysis come from a number of different fields and it is to be hoped that the continued attention of many investigators will increase the utility of the evaluative approach outlined in this paper.

REFERENCES

Agarwala, S. N.

1972 "A study of factors explaining variability in family planning performance in different states of India", paper presented at the All India Seminar on Family Planning Problems in India. Bombay, International Institute for Population Studies (mimeograph).

Alker, Howard R., Jr.

1969 "A typology of ecological fallacies", Quantitative Ecological Analysis in the Social Sciences, edited by Mattei Dogan and Stein Rokkan. Cambridge, M.I.T. Press, 69-86.

Berry, Brian J. L.
1971 "Problems of data organization and analytical methods in geography", Journal of the American Statistical Association, 66(335): 510-523.

Blalock, H. M. Jr.
1963 "Making causal inferences for unmeasured variables from correlations among indicators", The American Journal of Sociology, 59 (1): 53-62.
1967 "Path coefficients versus regression coefficients", The American Journal of Sociology, 72 (6): 675-676.

Chai, John J.
1971 "Correlated measurement errors and the least squares estimator of the regression coefficient", Journal of the American Statistical Association, 66(335): 478-483.

Chen, H. C. and Albert Hermalin
1970 "Factors affecting the performance of field workers in the Taiwan IUD program", Taiwan Population Studies Working Paper No. 7. Ann Arbor, Population Studies Center, University of Michigan (mimeographed).

Chung, Bom Mo *et al.*
1972 Psychological Perspectives: Family Planning in Korea. Korean Institute for Research in the Behavioral Sciences, Seoul, Korea, Hollym Corporation.

Coale, Ansley J.
1969 "The decline in fertility in Europe from the French Revolution to World War II", Fertility and Family Planning. Edited by S. Behrman, L. Corsa Jr. and R. Freedman. Ann Arbor, University of Michigan Press, 3-24.

Collver A., A. Speare and P. K. C. Liu
1967 "Local variations of fertility in Taiwan", Population Studies, 20 (3): 329-342.

Davis, K.
1967 "Population policy: Will current programmes succeed?", *Science*, 158 (November): 730-739.

Duncan, Otis Dudley
1964 "Residential areas and differential fertility", Eugenics Quarterly, 11 (June): 82-89.
1966 "Path Analysis: sociological examples", The American Journal of Sociology, 72 (1): 1-16.
1971 "Path analysis: sociological examples (addenda)", Causal Models in Social Sciences, edited by H. M. Blalock Jr. Chicago, Aldine-Atherton Inc., 136-138.
1972 "Unmeasured variables in linear models for panel analysis", Sociological Methodology. Edited by Herbert L. Costner. San Francisco: Jossey-Bass, Inc., 36-82.

Duncan, O. D., Ray P. Cuzzort and Beverly Duncan
1961 Statistical Geography: Problems in Analyzing Areal Data. Illinois: The Free Press of Glencoe.

Duncan, O. D., David Featherman and Beverly Duncan
1968a "Socioeconomic background and occupational achievement: extensions of a basic model", Final Report Project No. 5-0074 (EO-191). Office of Education, Bureau of Research, U.S. Dept. of Health, Education and Welfare.

Duncan, O. D., Archibald O. Haller and Alejandro Portes
1968b "Peer influences on aspirations: a reinterpretation", The American Journal of Sociology, 74 (2): 119-137.

Fawcett, James T.
1970 Psychology and Population: Behavioural Research Issues in Fertility and Family Planning. New York, The Population Council.

Freedman, Ronald
1967 "Applications of the behavioral sciences to family planning programs", Studies in Family Planning, 23 (October): 5-9.

Freedman, R. and John Y. Takeshita
1969 Family Planning in Taiwan. Princeton, Princeton University Press.

Friedlander, Stanley and Morris Silver
1967 "A quantitative study of the determinants of fertility behaviour", Demography, (4)1: 30-70.

Fruchter, Benjamin
1954 Introduction to Factor Analysis. New York, Van Nostrand.

Goldberg, Arthur S.
1966 "Discerning a causal pattern among data on voting behaviour", Causal Models in the Social Sciences, edited by H. M. Blalock Jr., Chicago, Aldine-Atherton Inc.: 33-48.

Gordon, Robert A.
1968 "Issues in multiple regression", The American Journal of Sociology, 73 (5): 592-616.

Hannan, Michael T.
1971 Aggregation and Disaggregation in Sociology. Lexington, D. C. Heath and Company.

Harmon, Harry H.
1967 Modern Factor Analysis. Revised edition, Chicago, University of Chicago Press.

Hauser, Robert M. and Arthur S. Goldberger
1971 "The treatment of unobservable variables in path analysis", Sociological Methodology. Edited by Herbert L. Costner, San Francisco, Jossey-Bass Inc.: 81-117.

Heer, David M. and John W. Boynton
1970 "A multivariate regression analysis of the differences in fertility of United States counties", Social Biology, 17 (3): 130-194.

Heise, David R.
1969 "Problems in path analysis and causal inference", Sociological Methodology. Edited by Edgar F. Borgatta. San Francisco, Jossey-Bass Inc., 38-73.

Hermalin, Albert I.
1971 "Taiwan: appraising the effect of a family planning program through areal analysis", Taiwan Population Studies Working Paper No. 14, Ann Arbor. Population Studies Center, University of Michigan (mimeographed).

Johnston, J.
1963 Econometric Methods. New York, McGraw-Hill Co., Inc.

Jöreskog, K. G. and D. N. Lawley
1968 "New methods in maximum likelihood factor analysis", The British Journal of Mathematical and Statistical Psychology, 21 (1): 85-96.

Kendall, Patricia L. and Paul F. Lazarsfeld
1955 "The relation between individual and group characteristics in 'The American Soldier' ", The Language of Social Research. Edited by Paul F. Lazarsfeld and Morris Rosenberg. New York, Free Press, 290-296.

Kitagawa, Evelyn
1964 "Standardized comparisons in population research", Demography, 1 (1): 296-315.

Land, Kenneth C.
1969 "Principles of path analysis", Sociological Methodology. Edited by Edgar F. Borgatta. San Francisco: Jossey-Bass Inc., 3-37.

Lazarsfeld, Paul
1958 "Evidence and inference in social research", Daedalus. 87 (4): 99-130.

Le Bras, Hervé
1971 "Géographie de la fécondité française depuis 1921", Population 6 (Nov-Déc.): 1093-1124.

Livi-Bacci, Massimo
1971 A Century of Portuguese Fertility. Princeton, Princeton University Press.

Lorimer, Frank
1959 "The development of demography", The Study of Population. Edited by P. Hauser and O. D. Duncan, Chicago, University of Chicago Press: 124-179.

Nerlove, Marc and S. James Press
1973 "Univariate and multivariate log-linear and logistic models", Rand Monograph R-1306-EDA/NIH, December.

Nerlove, Marc and T. Paul Schultz
1970 Love and Life Between the Censuses: a Model of Family Decision-Making in Puerto Rico, 1950-1960. (RH-6322-AID), Santa Monica, Rand.

Park, Chai Bin
1972 "Multivariate analysis of areal fertility in Honolulu", East-West Population Institute Working Paper No. 21. Honolulu, East-West Center.

Rhodes, Lewis
1971 "Socioeconomic correlates of fertility in the metropolis: relationship of individual and areal unit characteristics", Social Biology, 18 (3): 296-304.

Riew, John
1971 "The family planning programs in Korea—an analysis of achievement and cost", University Park, Pennsylvania State University (mimeographed).

Robinson, W. S.
1950 "Ecological correlations and the behaviour of individuals", American Sociological Review, 15 (June): 351-357.

Ross, John A. *et al.*
1972 "Findings from family planning research", Reports on Population/Family Planning, 12 (October). New York, The Population Council.

Rummel, R. J.
1970 Applied Factor Analysis. Evanston, Northwestern University Press.

Schoenberg, Ronald
1972a "Strategies for meaningful comparison", Sociological Methodology. Edited by Herbert L. Costner. San Francisco: Jossey-Bass Inc., 1-35.
1972b "Issues in multiple regression revisited", Paper presented at the 67th Annual Meeting of the American Sociological Association, August 28-31.

Schultz, T. Paul
1971a Evaluation of Population Policies: A Framework for Analysis and Its Application to Taiwan's Family Planning Program. (R-643-AID), Santa Monica, Rand.
1971b The Effectiveness of Population Policies: Alternative Methods of Statistical Inference, (P-4663), Santa Monica, Rand.
1973 "Explanation of birth rate changes over space and time: a study of Taiwan", Journal of Political Economy, 81 (2:2): 238-274.

Schultz, Theodore W. (Ed.)
1973 "New economic approaches to fertility" (Proceedings of a Conference, June 8-9, 1972), Journal of Political Economy, 81 (2-2).

Siegel, Paul and Robert Hodge
1968 "A causal approach to the study of measurement error", Methodology in Social Research, Edited by H. M. Blalock and A. Blalock. New York, McGraw-Hill Book Co., 28-59.

Simmons, Georges B.
1971 The Indian Investment in Family Planning. New York, The Population Council.

Taiwan Provincial Department of Health, Taiwan Population Studies Center
1966-68 Demographic Reference: Taiwan, Volumes IV, V and VII.

Taiwan Provincial Department of Health, Committee on Family Planning
1969 Family Planning Reference Book, Volume III.

Taiwan Provincial Government, Department of Civil Affairs
Taiwan Demographic Factbook. (Annually since 1961, with the exception of 1962.)

Taiwan Provincial Institute of Family Planning
1968 Family Planning Reference Book, Volume I.

Theil, H.
1954 Linear Aggregation of Economic Relations. Amsterdam, North Holland Publishing Company.

Thomas, Edwin N. and David L. Anderson
1965 "Additional comments on weighting values in correlation analysis of areal data", Annals of the Association of American Geographers, 55: 492-505.

Vig, O. P.
1972 "An application of path analysis to study variation in acceptance of the family planning programme in India, 1966-71", paper presented to the All India Seminar on Family Planning Problems in India, Bombay, International Institute for Population Studies (mimeographed).

Werts, Charles E., Karl G. Jöreskog and Robert L. Linn
1973 "Identification and estimation in path analysis with unmeasured variables", American Journal of Sociology, 78 (6): 1469-1484.

Zellner, Arnold
1962 "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Association, 57 (298): 348-368.

ACKNOWLEDGEMENTS

Analysis of many of the points discussed in this paper by Mr. Robert Johnson, as part of his graduate training program in population, was of assistance to the author. Research for this paper was supported in part by a grant from the National Institute of Child Health and Human Development.

Chapter 9
RETROSPECTIVE SURVEYS

by
K. Srinivasan *

I. INTRODUCTION

In recent years sample surveys on population have assumed an increasingly important role in the evaluation of the demographic impact of family planning programs. The data needed for analysis of the relationship between program factors and fertility change are not always obtainable from other common sources of demographic and family planning statistics, such as censuses, vital registration, population registers and family planning service statistics. Though each of these sources might provide data on one or more components, they do not have all the information needed for the same population for analysis of program effects at the individual couple or even the group level. Further, since measurement of change and inter-correlations at the micro-level are involved, there is a need to obtain information on the above factors and on changes in these over time for each couple or other sample unit. Sample surveys offer a unique opportunity for collecting such data. Developments in the field of sampling and estimation in the past few decades and the experiences gained in the application of survey techniques to other areas of the social sciences have made this an excellent way of evaluating family planning programs. In this chapter, a review is made of the applications and limitations of survey methodology in assessing the impact on fertility of family planning programs, focusing on the type, objectives, and scope of studies undertaken in the past, conceptual models underlying these studies, the variables and questionnaire construction, operational considerations, problems of sampling and non-sampling errors, techniques of analysis of data and priority areas for further research in the field.

* The author is currently working for the United Nations Economic and Social Commission for Asia and the Pacific. The views expressed in the chapter are the author's own, and do not necessarily reflect those of the United Nations or of the ESCAP secretariat.

II. AN OVERVIEW OF PURPOSES AND TYPES OF SURVEYS IN FERTILITY AND RELATED FIELDS

ORIGIN

The era of sample surveys for the study of fertility and related factors can be considered to have commenced with the study undertaken in 1941 on a sample of 1,444 relatively fecund white Protestant couples in the city of Indianapolis in the U.S.A. This study attempted to discover the social and psychological factors associated with fertility differences and to test some twenty-three hypotheses regarding their interrelationship (Kiser and Whelpton, 1953; Kiser, 1962).

Another important landmark was the survey undertaken by Lewis-Faning in 1946-1947 to answer specific questions posed by the Royal Commission on Population about the prevalence and effectiveness of contraceptive practices in Great Britain and their impact on fertility (Lewis-Faning, 1949). While the findings from the Indianapolis study gave impetus to further studies on fertility research, especially with regard to factors influencing fertility, the Lewis-Faning study showed the need for further understanding of contraceptive behavior and the effect of contraception on fertility patterns and trends. Since these two pioneering investigations, there has been a spate of studies relating to fertility and family planning in the developed and developing countries. In the developing countries a pioneering study in the application of survey methodology in estimating vital rates, contraceptive behavior and attitudes towards contraception and family size was undertaken in 1951-52 by the United Nations in Mysore State in India (United Nations, 1961). Studies in developed and developing countries have been reviewed by Berelson (1967).

The major problem in developing countries was not the explanation of fertility differentials, as in the developed countries, but the estimation of fertility levels and trends. The system of birth and death registration in many of these countries is of such poor quality that the data are of practically no value in the computation of fertility measures. Thus, sample surveys or sample registration schemes must be used for estimating fertility levels and trends.

With the increase in the number of family planning programs during the sixties, the interest and emphasis on fertility-related surveys have deepened and there has been a growing number of surveys on fertility-related topics. The scope and coverage of these surveys vary widely, however, depending on their specific objectives, the resources available, and other factors.

TYPES OF SURVEYS IN DEVELOPING COUNTRIES

If we consider only surveys which include questions on fertility (i.e., questions ranging beyond numbers of surviving children), those carried out so far in developing countries can be classified under three broad types: multi-purpose, demographic, and fertility/KAP, depending upon their major objective. The multi-purpose surveys are generally undertaken to gain an understanding and description of economic and social problems in a country and questions relating to population, such as household size and composition, births and deaths, form only a relatively small part of the questionnaire. In such surveys, obtaining demographic information is supplementary to the major objectives. National income and labor force surveys which are carried out in many countries on a periodic basis are illustrative of this type. The second category of surveys, those of the "demographic" type, have as their main objective the collection of reliable data on demographic factors, especially fertility and mortality rates, with a view to a better understanding of the nature and magnitude of population problems. The Population Growth Estimation Projects (PGE) (Eli Marks *et al.*, 1972) fall into this category. The third group of surveys, of a "fertility/KAP" type, are the primary concern of this chapter. These can be considered as descendants of the Lewis-Faning study and are generally preludes or adjuncts to family planning activities. Data are collected on fertility items and KAP variables: knowledge about the physiology of reproduction and methods of contraception; attitudes and values related to family size, pregnancy, family limitation in general, and the use of sterilization, abortion, and methods of contraception; practice of birth control.

The number of these surveys has increased dramatically in the past fifteen years and they have contributed variously to program assessment and improvement, to describing populations with regard to their fertility patterns, attitudes toward family size and family planning methods, and also to the methodology of survey research in general. If we include the small-scale KAP/fertility surveys carried out in selected communities by interested researchers or as part of a family planning program, the number of such studies would be around two thousand. Surveys having national or major geographic coverage are relatively few but have been increasing in recent years. An inventory of large-scale fertility and related surveys conducted since 1960 has recently been prepared by the US Bureau of the Census under the auspices of the World Fertility Survey (Duncan, 1973). It includes only those surveys having nationwide, regional, metropolitan, rural or urban coverage and while its main focus is the 200 such surveys conducted in developing countries since 1960, inform-

ation on 22 studies undertaken in developed countries is also included for illustrative purposes. The information given for each survey includes research design characteristics such as coverage, timing, survey type and method, sample size and questionnaire design. In a preliminary analysis these characteristics were cross-classified to illustrate different combinations of attributes, geographic variations, and broad time trends.

Tables 9.1, 9.2 and 9.3 below, which are taken from the above inventory, present data on the number of surveys undertaken since 1960 in developing countries classified by survey type and region, method and sponsoring organization.

Table 9.1. *Fertility and related surveys in developing countries by survey type and region (number and percentage)*

	Survey type			
Region	*Demographic*	*Fertility/KAP*	*Multipurpose*	*All types*
Asia	26 (39.4)	32 (48.5)	8 (12.1)	66 (100.0)
Africa	63 (75.0)	17 (20.2)	4 (4.8)	84 (100.0)
Latin America	7 (14.3)	38 (77.5)	4 (8.2)	49 (100.0)
Total	96 (48.2)	87 (43.7)	16 (8.1)	199 (a) (100.0)

Note: Figures in brackets represent percentages to row totals.
(*a*) Excluding those for which data on survey type were not available.
Source: Duncan, 1973.

Table 9.2. *Fertility and related surveys in developing countries by survey type and survey method (number and percentage)*

	Survey type			
Survey method	*Demographic*	*Fertility/KAP*	*Multipurpose*	*All types*
Single round retrospective	47 (36.4)	72 (55.8)	10 (7.8)	129 (100.0)
Multiround	31 (66.0)	10 (21.3)	6 (12.7)	47 (100.0)
Dual system	14 (100.0)	—	—	14 (100.0)
Total	92	82	16	190 (a)

Note: Figures in brackets represent percentages to row totals.
(*a*) Excluding those for which data on survey type and method were not available.
Source: Duncan, 1973.

Table 9.1 shows that over 90 per cent of the large-scale surveys conducted since 1960 in fertility and related fields in developing countries were of the demographic or fertility/KAP type and that equal numbers are found in each of these two categories. Surveys were most numerous in Africa and fewest in Latin America. In Africa, three out of four surveys were of the demographic type; in Latin America, fertility/KAP surveys were preponderant (over three-fourths); while in Asia, demographic and fertility/KAP type surveys were almost equally important. This is possibly due to the greater need for obtaining reliable estimates of vital rates in the countries of Africa, and to the importance of religious (Catholic) attitudes and values governing family planning and fertility behavior in Latin America.

Table 9.2 shows that about two-thirds of the fertility-related studies were single round retrospective surveys, and that of these, over half were of the fertility/KAP type and about one-third were of the demographic type. Two-thirds of the multiround surveys and all the dual system surveys were of the demographic type. Multiround and dual system are preferable to single-round retrospective enquiries when the aim is to obtain reliable data on vital rates.

Table 9.3. *Fertility and related surveys in developing countries by survey type and sponsoring organization (number and percentage)*

Sponsoring organization(s)	*Survey type*			
	Demographic	*Fertility/KAP*	*Multipurpose*	*All types*
International participation	9(28.1)	22 (68.8)	1 (3.1)	32 (100.0)
Foreign participation	33 (62.2)	18 (34.0)	2 (3.8)	53 (100.0)
Government	42 (55.3)	22 (28.9)	12 (15.8)	76 (100.0)
University	12 (32.4)	24 (64.9)	1 (2.7)	37 (100.0)
Total	96	86	16	198 (a) (100.0)

Note: Figures in brackets represent percentages to row totals.
(a) Excluding those for which data on survey type and sponsoring organization were not available.
Source: Duncan, 1973.

Table 9.3 shows that foreign or international participation has been involved in 85 out of the 198 surveys. About two-thirds of those involving international participation were of the fertility/KAP type and over 60 per cent of those involving participation of a foreign agency or government were of the demographic type. Among the 113 surveys sponsored solely by institutions within a country, 76 (55 per cent of them demographic) were conducted by governments and 37 (of which

65 per cent were fertility/KAP type) by universities. KAP studies appear to be quite popular with academic institutions, possibly because of the greater scope for research offered by such surveys.

Appendix I (Baum *et al.*, 1972) provides basic particulars on the fifteen fertility/KAP type surveys conducted or planned between 1968 and 1974 in Asia. These include summary information for each survey with regard to year, country and sponsoring organization, type of survey, approach to data collection, area and population coverage, sample size and stage of analysis and reporting. It is noteworthy that most of the surveys are single round retrospective surveys with a sample size of two to six thousand households and that in national surveys the emphasis is on fertility items rather than KAP data. Most of the KAP surveys conducted were in selected areas where intensive experimental action programs in family planning had been initiated.

The fact that we have put fertility and KAP surveys in the same category does not mean that the surveys in this group give equal emphasis to fertility and KAP items in their questionnaires. They include pure fertility surveys, such as the standard fertility surveys in India (Central Family Planning Institute, 1967) which seek to obtain data only on pregnancy or birth histories, closed and open intervals, periods of post-partum amenorrhea and lactation and current menstrual status of eligible women in addition to the usual questions on demographic particulars of members in the household. Practically no data of the KAP type are obtained in such studies. Similarly, there are surveys which are essentially of a KAP nature containing only very few items on fertility. During the past few years there has been a slow but steady fusion of these two types of surveys. Fertility surveys have begun to include a few KAP-type questions and similarly KAP surveys include questions on pregnancy or birth histories, justifying the common classification of such surveys under one heading.

III. OBJECTIVES AND APPLICATIONS OF FERTILITY/KAP SURVEYS

Fertility/KAP surveys have been undertaken in the past with a wide range of objectives and in recent years are finding increasing applications in the context of administration and evaluation of family planning programs. Until recently such surveys were confined to selected areas of a country where experimental family planning programs had been introduced. They were carried out under the auspices of academically oriented personnel from universities and research institutions. During the past few years the emphasis of these surveys has shifted towards national coverage and has even

at times aimed at international comparability, as with studies carried out as part of the World Fertility Survey. These broader surveys are usually carried out under the auspices of national governments or international organizations.

The objectives underlying the fertility/KAP surveys carried out in developing countries in the past have included one or more of the following:

(1) Description of the population with regard to demographic, socio-economic and KAP variables.
(2) Measuring current fertility levels and trends.
(3) Evaluating the impact of a family planning program through analysis of changes in KAP and fertility variables.
(4) Measuring private sector activities in family planning.
(5) Validating family planning service statistics.
(6) Generating support for a family planning program.
(7) Assisting program administration.
(8) Advancing social science.
(9) Institution building.

The first five objectives are aimed at obtaining information about the population and evaluating the impact of the family planning program. The last four are of a political and managerial nature and are intended to gain public support or funding for family planning programs or to make necessary administrative or institutional changes for improving their efficiency and effectiveness. Often, not all purposes can be covered adequately in any one survey.

The first objective, namely description of the population by significant demographic and KAP variables, has been the dominant part of most of the KAP surveys in developing countries and the findings have been widely published. Even when the objective is analytical, in the sense of identifying cause-effect relationships between variables, description is a necessary preliminary step. The data collected and analyzed with this objective in mind include tabulations on age-sex-marital status distributions of the population surveyed by selected socio-economic variables, such as income, occupation, etc., classification of married women in the reproductive ages by levels of knowledge of family planning methods, attitudes towards family size, family planning and program elements and current age-specific fertility levels. Current levels of fertility are estimated either on the basis of births in the previous twelve months or by indirect methods from the age distribution, number of children born and number living. The KAP surveys provide much needed data on the distribution of

population by significant demographic and socio-economic variables to serve as denominators for the computation of rates. The differentials in KAP and fertility variables observed in different socio-economic strata provide valuable insights for further investigation on the interrelationship between these variables. Table 9.4, drawn from a KAP survey conducted during 1962 in Taichung city in Taiwan, illustrates this type of differential analysis.

Table 9.4. *Proportion of couples with wives aged 20-39 who have ever used any or specific methods of family planning in Taichung Survey (1962) classified by wife's education*

	Percentage who have ever used contraception				
Wife's education	*Contra-ception*	*Abortion*	*Steriliza-tion*	*Any of three*	*Number of wives*
None	11	7	5	19	767
Primary school, not graduate	19	9	12	30	303
Primary school, graduate	29	13	9	38	842
Junior high school	46	20	13	57	320
Senior high school graduate or higher	66	22	12	72	211
Correlation rates	0.36	0.15	0.09	0.34	

Source: Freedman and Takeshita, 1969.

Table 9.5. *Relative influence of demographic and modernization variables*

	Proportion of total variance in birth control practice of couples attributable to:	
Type of contraception ever used	*Demographic variables (percentage)*	*Modernization variables after controlling for demographic variables (percentage)*
Contraception	37	63
Abortion	57	43
Sterilization	71	29
Any method	51	49

Source: Srikantan, 1967.

It is also possible to extend this type of analysis, through the application of multivariate techniques, to assess the relative influence of a number of demographic and modernization variables on family planning practice. For example, analysis of the data collected in the Taichung Survey in 1962 revealed the results shown in Table 9.5.

Fertility/KAP surveys have played a crucial role in bringing to public attention the desire for small families, favorable attitudes toward contraception and family planning and extent of practice of methods already prevalent even in the rural communities of developing countries. Table 9.6 illustrates this type of data with examples drawn from the West Malaysian survey (1966-1967). The demand for and public support to family planning programs is established by such data.

Similarly, after a program has been initiated, the extent and nature of its contact with the population, their perception of the program in general, and of program personnel and methods offered in particular,

Table 9.6. *Desire for additional children and approval of family planning among married women aged 25-34 and 35-44 in rural areas of Malaysia*

		Of those not currently using contraception		
Wife's desire for more children	*Number of women*	*Percent who expect to use in future*	*Percent who are interested in learning more about F. P.*	*Percent who approve of F. P.*
Wives 25-34				
Wants more children	328	43.8	67.4	67.4
Uncertain or does not know	17	—	—	—
Does not want more	147	54.8	75.5	72.6
Total	492	46.6	69.5	68.3
Wives 35-44				
Wants more children	150	26.5	46.7	55.4
Uncertain or does not know	17	—	—	—
Does not want more	185	39.7	62.6	68.5
Total	352	33.4	55.1	62.2

can be assessed and used in monitoring the program. The extent of changes in knowledge, attitudes and practice caused by the program can also be assessed if comparable data over time are available.

Measurement of changes occurring in the fertility/KAP variables of population over time is possible if the surveys are repeated on the same or independent samples at periodic intervals. In panel studies where repeated surveys of the same sample are made, it is also possible to identify causal mechanisms underlying the changes. Trend analysis by either of these two methods provides an important basis for understanding the dynamics of change in fertility/KAP variables, and when such comparisons are made before and after the implementation of a program, they are useful in the evaluation of its impact. Unfortunately, the extent to which findings from one survey are related to findings from other surveys conducted earlier has been limited. An illustration of the before and after comparisons of current practice of family planning in relation to experimental programs in different countries is given in Table 9.7.

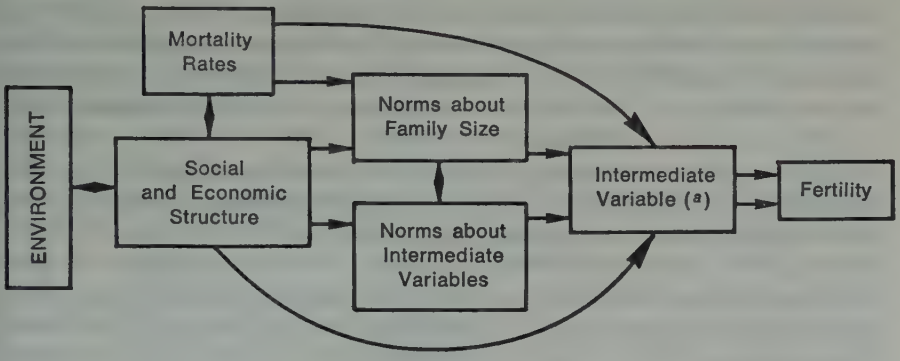
Table 9.7. *Percentage currently practicing family planning in selected before/after KAP surveys*

	Any method			*Wearing IUD*		
Study area and years	*Before*	*After*	*Increase*	*Before*	*After*	*Increase*
Korea (1964-1967)	9	20	11	<1	12	11
Koyang (1962-1964)	8	45	37	0	1	1
Kimpo (1962-1964)	11	20	9	0	7	7
Koyang County (1965-1966)	22	24	2	11	15	4
Sungdong Gu (1964-1966)	9	23	14	<1	10	10
Taiwan (1965-1967)	24	33	9	5	9	4
Taichung (1962-1963)	19	26	7	—	—	—
Thailand, Potharam (1964-1965)	4	29	25	—	—	—
West Pakistan (1966-1967)	—	—	—	2	4	2
India Hooghly District (April 1965-June 1965)	19	27	8	0	12	12

Source: Watson *et al.*, 1969.

IV. CONCEPTUAL SCHEMES

As mentioned in the previous section, sample surveys on fertility and family planning can be conducted purely for descriptive purposes or for analytical (explanatory or predictive) purposes or both. When the objectives of a study include an understanding of causal structure



(*a*) There is a limited set of variables through which any social or environmental variables must act to affect fertility. These are what Kingsley and Judith Davis call "intermediate variables." See Kingsley Davis and Judith Blake Davis, "Social Structure and Fertility: An Analysis Framework," Economic Development and Cultural Change, Vol. 4 (1956), pp. 211-235. They include:

1. Factors Affecting Exposure to Intercourse ("Intercourse Variables").

(*a*) Those governing the formation and dissolution of unions in the reproductive period.

(i) Age of entry into sexual unions.

(ii) Permanent celibacy: proportion of women never entering sexual unions.

(iii) Amount of reproductive period spent after or between unions.

(*a*) When unions are broken by divorce, separation, or desertion.

(*b*) When unions are broken by death of husband.

(*b*) Those governing the exposure to intercourse within unions.

(iv) Voluntary abstinence.

(v) Involuntary abstinence (from impotence, illness, unavoidable but temporary separations).

(vi) Coital frequency (excluding periods of abstinence).

2. Factors Affecting Exposure to Conception ("Conception Variables").

(vii) Fecundity or infecundity, as affected by involuntary causes.

(viii) Use or non-use of contraception.

(*a*) By mechanical and chemical means.

(*b*) By other means.

(ix) Fecundity or infecundity, as affected by voluntary causes (sterilization, subincision, medical treatment, etc.).

3. Factors affecting Gestation and Successful Parturition ("Gestation Variables").

(x) Foetal mortality from involuntary causes.

(xi) Foetal mortality from voluntary causes.

Any explanation of fertility trends and variations must include as a final step in the causal sequence changes in one or more of these "intermediate variables".

Source: Freedman (1961-62).

Fig. 9.1. *Schematic diagram of factors affecting fertility by Freedman including intermediate variables by Davis and Blake*

and interrelationship among variables, there is a need for specifying a conceptual scheme or model to facilitate formulation of hypotheses, selection of variables, and nature of analysis to be done. In a survey designed for evaluation of the demographic impact of a family planning program, there is a wide range of conceptual models in the literature from which a choice could be made with necessary modifications.

A conceptual model that has been very influential in fertility studies is the one formulated by Davis and Blake (1956), classifying a set of "intermediate variables" through which any change in fertility has to be mediated. It is an exhaustive classification in the sense that any change in fertility *must* be effected through changes in one or more of these intermediate variables. But it is not a mutually exclusive set in the sense that the intermediate variables are not independent of each other. The eleven intermediate variables were classified into three categories: (i) factors affecting exposure to intercourse (intercourse variables), (ii) factors affecting exposure to conception (conception variables), and (iii) factors affecting gestation and successful parturition. Freedman's elaboration of the Davis-Blake scheme within a broader sociological context, including a listing of the eleven intermediate variables, is given in Figure 9.1 (Freedman, 1961-62). It should be mentioned that the Davis-Blake classification lists only intermediate variables and does not attempt to explicitly postulate the nature of interrelationships among them, nor to explain fertility levels and change in terms of these variables. However, this classification assists in seeking appropriate variables to explain an observed fertility change. Sociological elaborations of the model have attempted to include the effects of norms about family size and about the intermediate variables themselves on each of these intermediate variables. Most of the fertility surveys undertaken in developing countries in the sixties have been based on such models. Illustrations are the models of Schultz (1971), Sirageldin (1972), Loewenthal and David (1972). In these models observed fertility (B) is considered to be made of two parts:

$$B = B_1 + B_2 \qquad (1)$$

where B_1 is number of wanted births and B_2, the number of unwanted births.

B_1, the number of wanted births, is further expressed as $B_{11} - B_{12}$ where B_{11} is the number of births desired for that period and B_{12} the number of desired births not realized due to various factors. Similarly B_2, the number of unwanted births, can be expressed as $B_{21} - B_{22}$ where B_{21} is the number of unwanted births that would have

happened in the period without the program and B_{22} the number of births averted by the program. Thus,

$$B = B_{11} - B_{12} + B_{21} - B_{22} \ldots \quad (2)$$

B_{11}, the number of desired births, is considered to be a function of utility or value of children, B_{12} being a function of fecundability, sterility, separation of spouses, etc., and B_{22} a function of the prevalence and effectiveness of contraceptive practice. For example, Loewenthal and David conceptualized that the utility of children (influencing B_{11}) and the ability of couples to control fertility (B_{12} and B_{22}) can be expressed as a function of sociological variables of the individual or family as follows:

(a) Utility of children is a function of family wealth, education of family members, family social status, the functional inclusiveness of the family, and the religion and religiosity of the family.
(b) Ability to control fertility is a function of education, status of women, access to family planning programs, and residual factors.

Schultz (1971) has considered a micro-model wherein the three environmental variables (education of children aged 12-19, extent of adult education and proportion of male labor force in agriculture) influence the number of surviving children parents want and the level of child mortality and family planning activities facilitate averting unwanted births. Though Schultz's model is individualistic, he uses aggregate areal data due to lack of equivalent data at the individual level. Schultz argues,

> "To infer how program activity influences birth rates, analysis of non-experimental data must simultaneously evaluate two sets of factors. First, environmental determinants of parents' family size goals and desired fertility must be identified and approximated by observed variables. Concurrently, family planning activities that reduce the number of ill-timed and unwanted births must be specified in as much detail and as comparably as possible across different target populations. Complete specification and accurate measurement of both sets of factors allow one to infer the independent contributions of each factor to the level of birth rates."

Under different functional specifications of the model Schultz has estimated the average and marginal effects of the program on the birth rate and the contribution of various forms of personnel input into the program.

A similar macro-model has been developed by Sirageldin (1972) for the evaluation of the family planning program in Pakistan. This model is a feedback model in the sense that the direction of relationships among fertility levels, social setting and program input variables in a population is considered both ways and not one-way, as is assumed in the Schultz model. The dependent variables used by Sirageldin in the analysis are fertility and contraceptive knowledge and practice. The explanatory variables are classified into four sets as follows:

(a) Background factors: age, rural/urban background, education, past mobility.
(b) Situational factors: general economic situation, current place of residence, ecological factors.
(c) General value system: attitudes towards hard work, ambition, planning ahead.
(d) Program input and facilities.

Fawcett (1970) argues that psychological variables have been given inadequate attention in the explanation of fertility differentials or contraceptive behavior and suggests adoption of a psychological functional approach to the study of attitudes and behavior, postulating that attitudes and behavior patterns are developed and maintained on the basis of their utility for regulating the interaction between personality and development. He also suggests that various psychological theories be applied in the design of fertility/KAP studies, not only to explain attitudinal and fertility differentials of couples but also to develop more effective communication strategies. Easterlin (1969) has suggested a model in which child-bearing choices are balanced in terms of utility against other kinds of choices on the basis of an economic-decision approach to fertility. Not only the sociological or psychological value of children but also the perceived costs (direct and indirect) of child-bearing and rearing should enter into the analysis.

A comprehensive model to explain fertility differentials which incorporates sociological, psychological and economic variables simultaneously has yet to be developed. An attempt at synthesizing the sociological and economic models has recently been made by Easterlin (1972). He considers the determinants of fertility as working through one or more of the following:

(1) desired family size (C_d), the number of surviving children parents would want if fertility regulation were costless,

(2) the potential supply of children (C_n), the number of surviving children parents would have in an unregulated fertility regime, and
(3) the costs of fertility regulation, including both subjective (psychic) costs and objective costs: the time and money required to learn about and use specific techniques.

In addition to the above, in recent years there have been notable contributions in the development of mathematical or probabilistic models of human reproduction. In this approach fertility is considered a random variable stochastically dependent on such factors as duration of marriage, fecundability, incidence of foetal wastage, postpartum amenorrhea, and incidence of sterility. Perrin and Sheps (1964) considered human reproduction as a zero order renewable Markov process dependent on the above variables. By assuming certain distributional forms for each factor, they derived explicit expressions for asymptotic fertility rates, mean birth intervals, etc. Family planning programs are assumed to modify fertility by changing the fecundability levels of the population. Sheps (1967) has outlined the utility of such models in the evaluation of effectiveness of contraceptive methods on population policies.

V. HYPOTHESES

Formulation of hypotheses is a fundamental step to be taken before the commencement of any survey which has an explanatory or predictive purpose and will determine the direction and scope of the survey. Their nature depends upon the objectives of the study and the conceptual model assumed to underly the interrelationship between the dependent and explanatory variables. A particular study objective can lead to a variety of hypotheses, depending upon the conceptual scheme and the assumed functional relationship between variables. They may be couched in such specific statistical parlance as "null hypotheses" or worded generally in the form of questions for which information is sought through the surveys. In most of the KAP/fertility surveys the hypotheses are framed in terms of general questions or specific patterns and trends expected among variables and not as "null hypotheses". Trends and directions of differences between different groups, rather than the statistical significance of differences, are the object of enquiry. For example, the Lewis-Faning study (undertaken in 1946-47) attempted to answer the following ten questions with regard to the contraceptive and fertility behavior of the population of Great Britain:

(1) How extensively is birth control practiced?
(2) In what proportions are the different methods of birth control practiced?
(3) Are there important differences between different social groups in the extent of the practice of birth control or in the choice of method?
(4) To what extent is birth control, as practiced, effective?
(5) What is the extent of involuntary infertility?
(6) Does the practice of birth control affect the power to reproduce?
(7) How important is abortion as a method of birth prevention?
(8) What is the proportion of "unplanned pregnancies"?
(9) What is the proportion of "unwanted children"?
(10) What are the chief reasons given for using birth control?

This study can be termed a descriptive study in the sense that it attempted to obtain information on certain areas relating to contraceptive practice and fertility behavior without any predetermined design to test interrelationships, though such analysis was carried out after the collection of data. Testing the statistical significance of differences was of subsidiary interest. Each of the ten questions quoted above has to be defined in operational terms in order to facilitate collection of data through a schedule or questionnaire.

Similarly, the study undertaken in Mysore State in India under the joint auspices of the Government of India and the United Nations (1961) was an experiment in methods of obtaining information on vital rates, on factors influencing them and on their social and economic consequences in rural and urban areas. It was a methodological experiment in the use of sample surveys for estimating population size, fertility and family planning parameters. The study that was undertaken in Taichung by Freedman and Takeshita (1969) attempted to provide answers to the following questions:

(1) What were the existing patterns of fertility and family planning before the organized program?
(2) How effective was the experimental family planning program?
(3) What happens after acceptance?

Each of the above three questions was broken down into a few operational questions requiring specific information from the survey. For example, the operational questions on (3) above were:

(a) What determines, for example, how long an IUD is kept in place and why its use is terminated?

(b) Once a termination occurs, what is the probability that birth control will be continued by other methods; that is, what is the history of the user rather than of the device?
(c) How many births are averted or prevented by the program, either by use of the initial device accepted or by something that is adopted afterwards?

Each of the above questions was broken down into a series of hypotheses with operational definitions of the terms involved.

Many social scientists argue that studies, whether involving surveys or not, will not reach particularly useful conclusions unless they have an initial set of hypotheses and an organized conceptual structure; there should be clear purposes beyond asking some interesting questions. It is also argued that fertility/KAP surveys must be designed with specific program goals in mind in order to achieve their full descriptive and analytical potential. Sirageldin (1972) argues that any survey undertaken for evaluation of the impact of a family planning program is attempting to answer one or more of the following three sets of questions, which may or may not be explicitly stated:

> "The *first* is related to the setting of a target reduction in the rate of population growth in a given period and to establishing how much a society can or should spend on family planning programs as opposed to other alternative investments..." (cost-benefit orientation).
> "The *second* set of questions is more substantive and related to program design and operations in finding the most efficient ways to achieve program objectives, given budget, personnel and time constraints..."
> "The *third* set of questions is related to the evaluation of program activities and achievements... It includes evaluation of program efficiency (e.g., cost-effectiveness), and the evaluation of fertility decline..."

VI. VARIABLES AND QUESTIONNAIRE

A. *VARIABLES*

The variables on which information has to be collected in any KAP/fertility survey undertaken for the evaluation of program impact can be grouped under four headings: fertility, contraception, socio-economic factors and program variables. These are briefly discussed below.

FERTILITY. Measurements of fertility can be based on live births, pregnancies and birth intervals. Further, they may be on the basis of cohort or period data. When data on detailed pregnancy histories are

obtained from each woman included in the sample, it is theoretically possible to construct various fertility measures. However, large errors are found to exist in the pregnancy history data obtained in developing countries. In the analysis of fertility change occurring over a short period of time not all measures are equally sensitive. For example, short-term changes in age-specific fertility rates could occur because of changes in the proportion of women married in that age group and in the proportion of married women in the susceptible condition (menstruating) in the previous year. Again, computation of fertility rates needs data on denominators and errors in the numerator and denominator could act independently to cause sizeable errors in estimates. Fertility indices based on birth intervals, such as average closed interval (interval of time between two successive births) and average open interval (interval of time from the last birth to date of survey) do not need separate denominator data, and the effects of contraception in delaying births are more readily reflected in interval data than in birth data. Analytical investigations have revealed that the mean open birth interval taken as an average over a group of married women in the reproductive ages can be used as a sensitive indicator of current fertility, and when fertility declines because of recent contraceptive practices among couples, the mean open birth interval can be expected to increase (Srinivasan, 1971). Empirical studies carried out in Taiwan and Korea reveal that differences in the open birth intervals explain most of the fertility differentials among groups and that when fertility declines, it is reflected in the lengthening of open birth intervals. When couples resort to birth limitation practices rather than spacing of children (as is found to be the case in most developing countries), the effect on fertility can be detected even in the short run in the lengthening of the mean open interval taken at successive points in time. Since the information needed is only the date of birth of the last child for every woman, the reliability of data on this interval can be expected to be quite good, even in developing countries. Information on this variable is available in almost all KAP/fertility surveys. Some of the limitations and consequent caution needed in the use of birth interval data due to the presence of certain types of biases, different approaches that could be adopted in the ascertainment of intervals and difficulties in separating period and cohort effects have been pointed out in recent years by Sheps and Menken (1973) and Venkatacharya (1969).

CONTRACEPTION AND FAMILY SIZE ATTITUDES. Questions on attitudes toward family size and KAP of family planning have formed the bulk of questionnaires for KAP/fertility surveys. Informa-

tion on knowledge, attitudes and practice of various methods of contraception and attitudes towards family size have been collected with varying emphasis in different surveys. For the measurement of impact on fertility, information on knowledge and attitudinal variables are only peripheral, but they help to measure the program impact at an intermediate level. There is an implicit assumption that changes in knowledge and attitudes are necessary precursors for contraceptive practice and that increased knowledge will lead to favorable attitudes, which in turn will lead to effective contraceptive practice and fertility decline. This simplistic assumption may be questioned (Kingsley Davis, 1967). Collection of data on knowledge and attitude variables will not only help evaluation of the program at an intermediate impact level but may also provide clues for reorganizing the communications and service strategy of the program in order to make it more effective. They thus serve directive as well as evaluative roles. With regard to information on practice, the ultimate goal is to obtain a detailed contraceptive history for each woman in the reproductive age group over her entire married life, including periods of irregular or non-use, accidental pregnancies and induced abortions. Such a contraceptive history can then be related to the pregnancy history to analyze the nature of the interrelationship.

The IUSSP list (United Nations, 1970) included family size and KAP variables under the following four headings:

(1) Attitudes toward family size and child spacing.
(2) Attitudes toward family limitation.
(3) Knowledge and communication about family limitation.
(4) Past and expected practice of family limitation.

The following variables were recommended in its core-list.

(a) *Attitudes towards family size and child spacing*

1. Respondent's current desire for additional children.
2. Preferred number of children, if the respondent could choose number.
3. Respondent's attitudes about family size for the general population.
4. Attitude toward last pregnancy.
5. Desires and plans for family growth.

(b) *Attitudes towards family limitation*

1. Respondent's attitude toward the general idea of family limitation.
2. What are her reasons for being for or against family planning?

3. Is her husband for or against the idea of family limitation?
4. Respondent's attitude toward induced abortion.

(c) *Knowledge and communication about family limitation*

1. Does respondent know any specific ways to postpone or avoid pregnancy? (Yes/No)
2. If respondent knows about contraception, knowledge of methods and how to use them.
3. Respondent's perception of whether contraception is used by her relatives, friends, neighbors, the general population and by significant elites in the society.
*4. Respondent's exposure to mass media.
*5. Has the respondent ever heard about family planning from any of these media?
6. Would the respondent like to learn more about family planning?
*7. Awareness and knowledge about some salient features of organized family limitation programs.

*relates to program exposure.

(d) *Past and expected practice of family limitation*

1. Has the respondent ever done anything to delay or prevent pregnancy?
2. If she has ever used a method, date of commencement, types of methods used, current use, and source of information for the last method used.
3. If respondent is not currently practicing contraception, future intention on use of methods.
4. Has the respondent ever had an induced abortion? If so, the number and order of pregnancies terminated.
5. Future intention on use of induced abortion.
6. If sterilized, date and type of sterilization operation undergone.

The concepts of expected, desired and ideal family size have been widely used in KAP/fertility surveys in developing countries. It has been found that there is a high correlation between these variables and acceptance and effective use of contraception.

PROGRAM VARIABLES. Information on program variables is needed for relating program input to demographic impact. However, this is the weakest set of variables in most KAP surveys. If the

evaluation is done at the macro or areal level, information on such factors as number of clinics, number of medical and field personnel per thousand population, proportion of trained to untrained workers, program strategy, such as mobile versus static clinics, postpartum program, number of home visits, can be collected for each areal unit and related to levels of contraception and fertility. This type of macro analysis relating program input to contraception and fertility behavior has been attempted by Schultz (1971) and Hermalin (1968). However, when analysis is attempted at the micro-level, i.e., the individual or the family, as is the case when a sample survey of households or eligible women is undertaken, the information needed is on factors such as nature and frequency of exposure of the woman or her husband to program personnel or services, the sociological and psychological acceptability of services rendered as perceived by the couples, follow-up services after acceptance of contraception, etc. Information should also be collected for differentiating the direct from the indirect effects of the program.

For example, couples could make use of clinic services offered within the program (direct effect) or from non-program sources, such as private physicians or drug stores, having been motivated to do so by the program (indirect effect). The separation of these two effects is possible only if data are collected on source of information on family planning and motivation for adoption of contraception for each couple and the places from which contraceptive supplies and services are obtained. Information is also needed on the extent of substitution of pre-program or traditional methods of contraception by program methods, and within program methods between program and non-program sources. The importance of data of this type for estimation of program impact has been brought out by Potter (1970). Nortman (1972) has compiled comparative data on the proportions of married women using different methods of contraception from program and private sources for different countries with national family planning programs where such data have been collected through surveys.

SOCIO-ECONOMIC AND DEMOGRAPHIC VARIABLES. Socio-economic and demographic variables are used as control variables in analysis relating program variables to contraceptive or fertility behavior. A conceptual model usually underlies the selection of such variables (see Section III, page 10). From the experience of past surveys it is clear that sociological factors such as education and employment of women and rural/urban status are correlated with fertility behavior even when the effects of other factors are controlled. However, the effect of any one variable on fertility

is small and varies from culture to culture. A summary of the findings from various surveys and studies with regard to socio-economic correlates of fertility can be had from Loewenthal and David (1972). According to the authors,

> "In this review, we have organized the potential causes of family fertility into four categories: the stratification variables, the division of labor between family and society, the division of labor between the sexes, and religious and value system variables. Of these four, the first is probably the most frequently studied and from a policy stand-point may be the most important to understand. Our review suggests, however, that in this area, as in the others, our understanding is still limited. We have some idea as to which variables correlate with fertility and the direction of these correlations, but how all these variables fit into a *causal system* is not well understood at this time. Also understood in only the most rudimentary way, in most instances, is how 'gross' social and economic indicators operate through social-psychological intervening variables to actually affect fertility."

The IUSSP sub-committee has recommended the following variables in its core-list.

(1) Educational attainment of wife and husband.
(2) Rural-urban status of current residence.
(3) Rural-urban background of husband and wife.
(4) Details of current and past economic activity of wife.
(5) Current economic activity of husband.
(6) Household composition (age, sex, marital status, and relationship to head of household).
(7) Religion and religiosity of husband and wife.
(8) Other cultural characteristics, such as ethnic and language classification.

In addition to the above, data on demographic characteristics and identification particulars of the members of the sample household, such as name, age, sex, marital status and relationship to the head of household are obtained to provide denominator data for the computation of rates and the selection of a sample of married women in the reproductive ages for detailed KAP/fertility interviews.

The questionnaires used so far in KAP/fertility surveys have been quite lengthy and have required between half an hour and two hours for an interview. For example, the questionnaire used in the West Malaysian Family Survey (1966-67) was 46 pages long, consisted of 319 items and took an average of more than two hours for an interview. A comparative analysis of the contents of the model questionnaires of the IUSSP and Population Council (1970) is provided in Appendix 2.

B. *QUESTIONNAIRE*

The IUSSP Model Questionnaire is 32 pages in length and the average interview time is expected to be over an hour. It contains 114 questions (233 including parts), though not all are applicable to every woman. The Population Council Model (1970) is 20 pages in length, contains 48 questions (132 including parts) and is expected to take over half an hour to complete.

The greater emphasis on a detailed pregnancy history and the probing questions on attitudes and knowledge of the IUSSP questionnaire make it the more useful one for fertility studies undertaken for scientific purposes, while the Population Council questionnaire will be more useful for program evaluation purposes. However, both questionnaires are generally weak in program-related variables, for example, the intensity of exposure of women to various program components, such as mass education, contact by family planning personnel, knowledge on how to use contraceptive methods, etc. Even on contraceptive history neither obtains complete information, the IUSSP model having nine questions on past use, while the Population Council has only one question on this topic.

Both questionnaires devote more than one-third of the total number of questions to obtaining information on socio-economic and demographic characteristics, 83 out of 233 in the former and 59 out of 132 in the latter. These variables are intended to serve essentially as control or explanatory variables in the study of differential fertility. From this point of view the questionnaires are better fitted for scientific purposes, in seeking a better understanding of sociological factors affecting fertility, rather than program evaluation purposes.

VII. STUDY DESIGN

STUDY UNIT AND TIME DIMENSION

When a survey is undertaken for the purpose of evaluating the KAP/fertility impact of a program without any predetermined design of experimentation, no unequivocally valid conclusions can be drawn on the cause-effect relationship between the program and changes in the KAP/fertility variables observed in the population. The program effects on fertility computed on the basis of analysis of data collected through retrospective surveys in such conditions can at best be only approximations of results achieved from experimental designs. This is because the principles of randomization in the application of program elements among population study units and the replications which are

fundamental to obtaining statistically valid conclusions cannot be fulfilled in an ongoing program. Only designed experimentation can lead to such results. The generalization of findings from retrospective surveys has to be done with caution. The data collected through surveys are, however, valuable for descriptive purposes and for testing hypotheses based on a conceptual model.

The problems and procedures involved in designing a KAP/ fertility survey for the purposes of family planning program evaluation are basically similar to those met with in the evalution of any social action program. The essential steps in the design of such a survey have been outlined in many standard textbooks on research methodology (e.g., Jahoda, Deutsch and Cook, 1951).

The unit of study for the evaluation of a family planning program could be the adult population, women in the reproductive age group, couples, families, or the community as a whole. Most of the KAP/ fertility surveys have concentrated on married women in the reproductive age group. A few have collected additional data on fertility items from ever-married women in the age group 15-50 and on KAP variables from a sub-sample of the husbands of women included in the sample. The IUSSP Committee recommends this approach. In a few studies, such as those conducted by Hill, Stycos and Back (1959) in Puerto Rico, the unit of study has been the nuclear family. If the objectives of a survey are not only to assess the impact of a program but also to help identify effective means of communicating and motivating couples to adopt family planning in order to improve its impact, then the unit of study should be one towards which educational and service activities of the program could be directed. Interviewing husbands and key decision-makers within and outside the family assists in this purpose. Whole communities or whole population groups as study units have also been used in anthropological investigations in a few countries.

Any analysis of program impact has a time dimension built into it requiring assumptions about the time gap between various inputs into the program and expected outputs. The impact of programs on acceptance, continuity of use, births averted and fertility levels occurs with a time lag, which may vary from country to country and over time within a given country. Hence, in order to measure trends of program impact, repeated surveys over time are needed. A single round retrospective survey will not be able to measure with reasonable accuracy the trends and changes in all these variables over time. Multiround surveys using independent samples at each round, or panel studies wherein the same sample is interviewed repeatedly at successive rounds, are usually adopted for the study of such trends.

The special advantages of panel studies are their value for time series analysis, for the study of errors of recall lapse and for their built-in ability to improve the reliability and accuracy of data on sensitive topics, such as pregnancy status, foetal loss, amenorrhea, etc. Since the same woman or household is visited at periodic intervals of time, information obtained in previous visits can be utilized to check and improve the quality of data of succeeding visits. For example, when a woman says that she is three months pregnant in the first interview and if the second interview is made six months later, the woman can be questioned in relation to her earlier condition and her current pregnancy status, and the nature and timing of termination of the pregnancy (if it had terminated in the sixth-month period) can be more accurately and reliably ascertained. If a live birth has occurred two months before the second round, the data on the month of gestation at the time of the first round could be corrected on the basis of this information. In a panel study in India where a sample of married women in the reproductive ages were interviewed at successive six-monthly intervals of time, it was found that there were substantial differences in the proportions of women at different months of pregnancy as obtained at the time of the first survey, and corrected by the information obtained in the revisit made six months later. Table 9.8 illustrates this point.

The table reveals that by adopting a questionnaire method, only

Table 9.8. *Percentage of pregnancies of less than eight weeks and over eight weeks gestation detected at each round (when information on subsequent round is considered for correction)*

Round of survey	*Period of gestation at the time of visit*	
	Eight weeks or more	*Less than eight weeks*
I	67.4	18.0
II	72.7	15.8
III	51.4	11.1
IV	54.6	12.3
V	74.0	11.7
VI	71.0	19.7
VII	65.3	15.2
Average	65.4	14.8

Source: Srinivasan *et al.*, 1969.

14.8 per cent of the pregnancies of less than eight weeks gestation and 65.4 per cent of those of over eight weeks could be detected. Data on the pregnancy status of women at any given point in time are significantly improved by using the information on pregnancy status gained from follow-up visits.

In panel studies the quality of data is sequentially improved in successive rounds and the timing of occurrence of various events can be better charted, permitting a variety of statistical analyses. The difficulty with this approach, however, is that the data may become more and more biased at every stage and may become less representative of the population. For example, when the same panel of women in reproductive ages is followed over a five-year period, at the end of the period the mean age of the panel will be five years greater than its mean age at the beginning of the study, and though the panel was representative of the population of women in the reproductive ages at the beginning of the study, this representativeness is slowly lost; and at the end of the study, unless necessary adjustments in the estimation procedures are made, there will be built-in biases in the findings. Also, frequent interviewing may make the women more curious and interested in the field of investigation and cause attitudinal and behavioral changes which are not representative of the phenomena occurring in the whole population.

SAMPLING

The problems connected with sampling for fertility and KAP surveys are essentially the same as for surveys in other fields. The basic objective of selection of any sample is to achieve maximum precision for the needed estimates at unit cost. The forumlae for the standard errors of some of the important statistics based on a sample usually rest on simplistic assumptions about the nature of the distribution of characteristics.

An illustration of the minimum sample size needed for detecting differences of various magnitudes in general fertility rates and age specific fertility rates for age-group 30-34 is provided in Table 9.9. The computations are based on the assumptions that the general fertility rate (f_1) and age specific fertility rate (f_2) follow a binomial distribution; a simple random sampling procedure is adopted to obtain estimates, and the population parameter values are: average household size 5.4, 22 per cent of the population are women in the age group 15-44, 3 per cent of the total population are women in the age group 30-34, the general fertility rate is 0.195 and the age-specific fertility rate for ages 30-34 is 0.3. The parametric values assumed

are close to those observed in many developing countries. The sample size is based on a simple test of significance of the difference of two sample proportions using confidence intervals of two times the standard error around the mean value.

Table 9.9. *Sample size needed for detecting various percentage changes in fertility rates*

Percentage change in		*Minimum sample size needed to detect change*			
General fertility rate (15-44) (f_1)	*Age specific rate for age 30-34 (f_2)*	*Total population*	*Number of households*	*Number of women 15-44*	*Number of women 30-34*
4	8	100,000	18,500	22,000	3,000
12	25	10,000	1,850	2,200	300
17	35	5,000	925	1,100	150
22	46	3,000	550	660	90
27	56	2,000	375	440	60
39	79	1,000	175	220	30

Source: Adapted from Seltzer, 1970.

From table 9.9 it can be seen that in order to detect a 4 per cent change in the general fertility rate or an 8 per cent change in the age specific fertility rate of age group 30-34 with statistical significance at 95 per cent confidence level, a minimum sample of 100,000 people or 18,500 households are to be interviewed. When the change in the general fertility rate is of the order of 12 per cent or more, a sample of 10,000, or 1,850 households would be needed. This table reveals that in order to detect a small change in fertility (less than 5 per cent in the general fertility rate) a fairly large sample of about 15,000-20,000 households is needed. Only when large changes in fertility rates are expected to occur in the population would the smaller samples which are usually used in KAP/fertility surveys suffice.

From Appendix I, it can be seen that most of the fertility/KAP surveys have a sample size of two to four thousand households. With this range of sample size there would need to be a change in the general fertility rate of over 10 per cent for it to be detected from the sample with statistical significance.

Fixing the sample size is relatively easier in cases where only

one parameter, such as the crude birth rate or general fertility rate, is to be estimated. When estimation and analysis of interrelationships of many variables is involved, as in a KAP/fertility survey, the problem becomes complicated, even theoretically. Ideally, the sample size needed for estimation of each of the parameters or correlation coefficients should be determined for realizing a specified level of accuracy, and the maximum in this set of samples should be considered as the desired sample size for the survey. In many situations where one or two variables with very low frequency of occurrence in the population are included in the survey, this will necessitate samples so large as to be operationally unfeasible. Further, the preliminary estimates of mean and variance of a variable which are needed for determination of a sample size may not be available for many of the variables included in a KAP survey, and there may be a need to undertake a pilot survey to obtain them. Very few KAP/fertility surveys have undertaken pilot surveys for this purpose. However, since the occurrence of most of the variables included in a KAP survey, such as proportion having knowledge of family planning methods, having favorable attitudes, etc., are found to be relatively higher than fertility rates, the latter generally form the basis for estimation of sample size. In practically none of the surveys with national coverage have KAP variables been used as a basis for determining the sample size. On the contrary, the sample sizes are based essentially on the financial and other resources available for the survey and the maximum number of interviews that could be undertaken with the particular schedule of interview, distribution of sample units in the population, quality of interviewers available, etc. Even when fertility indices form the basis for theoretical determination of sample size, the theoretical basis for the assumption of a binomial law for the distribution of fertility in a population can be questioned. More complicated distributions based on stochastic processes are being tried for describing human fertility, but further research is warranted in this field before they can be applied for estimation of sample size (Perrin and Sheps, 1964).

The choice of a sampling method is dependent on a variety of factors, such as the extent of heterogeneity of the population, the stratification variables that can be used, data availability on the variables in the sampling frame, administrative convenience for organizing the survey, etc. From Appendix 1 it can be seen that a combination of cluster and stratified sampling has been adopted in most of the fertility/KAP surveys. Adoption of cluster sampling is operationally convenient when the geographical coverage of the study is widespread. When a whole village or block of a town is included in the sample, it is possible to conduct the interview at lower financial and

other costs than in a situation where the sample households are distributed in different villages or towns. This explains the popularity of cluster sampling in most of the large-scale sample surveys. However, it has been found from the experience of past fertility/KAP surveys, such as the West Malaysian Family Planning Survey in 1966-67, that higher sociological and cultural homogeneity within a village or an urban block may violate the basic criterion for choosing cluster sampling, that clusters should be highly heterogeneous within and homogeneous without.

VIII. ANALYSIS OF DATA

The techniques of analysis of data collected through KAP/fertility surveys vary widely in their complexity, ranging from the formation of simple frequency distributions and two-way tables on the one hand to the application of advanced statistical methods in multivariate analysis on the other. The process of analysis proceeds from the simple to the complex and the level of sophistication adopted in any particular study depends on the objectives of the survey, the study design and the technical and financial resources available for data analysis. Poor quality data do not justify application of sensitive analytical techniques, especially for the study of secular changes and interrelationships between variables.

When the objectives of the study are essentially descriptive, simple statistical analysis, such as frequency distributions, computation of averages and percentages, and two-way tabulations (bi-variate analysis) would meet the purpose. The bi-variate analysis is intended to study the linkage between a dependent variable such as fertility rates or acceptance rates and an independent variable, such as exposure to the program, availability of services, etc. One of the simplest ways of evaluating a program is to compare the acceptance rates or fertility rates among areas exposed to alternative approaches, between sub-components within a program, between treatment and control groups and so on. Comparisons can also be done among areas served by different categories of family planning workers, by different systems of delivery services and through varying emphasis on different methods of contraception. A major problem with such bi-variate analysis is that a sufficient number of cases may not be available for each cell to provide statistically reliable results.

Comparisons of levels and changes in the dependent variable classified by various categories of an independent variable are valuable because they are easily made and understood. However, if the

nature of the relationship between the two variables has to be quantified, this result should be used for bi-variate analysis, such as correlation and regression analysis. Such analysis is possible only when both the independent and dependent variables are in interval scales. When the independent variable is in a nominal or ordinal scale, resort can be had to the technique of analysis of variance, a powerful statistical procedure which attempts to partition the total variance of the dependent variable in two parts, that due to between-the-categories variation of the independent variable and that due to within-category variation. When the dependent variable is measured on a continuous scale and the independent variable as a nominal or ordinal scale, even correlation or regression analysis could be adopted using dummy variables. Correlation and regression coefficients computed from bi-variate analysis should not be construed as measures of causal relationships but merely suggestive of linkage or association.

Extensions of bi-variate analysis (comparisons, correlation/regression analysis, and analysis of variance) can be undertaken taking into account more than one independent variable simultaneously. The most common forms of multi-variate analysis are multiple regression analysis, analysis of covariance, multiple classification analysis and path analysis. The problem of studying the nature of interrelationships between a set of independent variables and a dependent variable can be considered from two angles: one, how well all the independent variables together, or each variable considered separately (taking into account or ignoring the other independent variables) explain the variation in the dependent variable; second, how well one can predict the score on the dependent variable for a person for whom the scores on the independent variables are given. The techniques of analysis of variance, covariance and correlation and multiple regression analysis attempt to answer both these questions.

Two important conditions for the application of traditional methods of analysis of variance of many independent variables on a dependent variable are that the independent variables are uncorrelated among themselves and that the interaction effects are not significant. The absence of correlation among the independent variables is usually achieved by designing the experiment in such a way that there are equal numbers of observations in each cell. Where there is an unequal number of observations in different cells, the analysis becomes extremely complicated. On the second condition, when the interaction effects are significant, analysis of variance fails to give meaning to the main effect of each independent variable. On the other hand, multiple regression analysis is capable of meeting

the above two defects of analysis of variance (intercorrelation between predicted variables and interaction effects) but requires that all independent variables and dependent variables be measured as interval scales.

Since most of the variables on which data are collected through KAP/fertility surveys, such as knowledge of contraception, attitudes towards program personnel and contraception, methods and nature of exposure to the program, are nominal or ordinal scales, application of traditional multiple regression analysis is not possible. In addition, many of the independent variables used are intercorrelated and have significant interaction effects on the dependent variable, thus violating the conditions for the application of the traditional technique of analysis of variance. For example, the effect of formal education of couples on contraceptive practice can be expected to be influenced by the availability of family planning services and the proximity of clinics to the residence of the couples. Thus, both analysis of variance and classical regression analysis cannot be applied to most KAP data.

Multiple Classification Analysis (MCA) overcomes many of the drawbacks inherent in the two traditional analytical procedures discussed above. MCA is a technique for examining the interrelationships between several predictor variables and a dependent variable in the context of an additive model. The only essential conditions are that the dependent variable be measured as an interval scale and that it be capable of being expressed as an additive function of the predictor variables. It allows for weak measurement (nominal or ordinal scale) on the predictor variables, intercorrelation among them and non-linear relationships. The technique is used to obtain (i) the net effect of a predictor on the dependent variable (in the deviation of the category mean from the grand mean after adjusting for the effect of other predictors) and (ii) the proportion of variance in the dependent variable explained by a group of predictors. A computer program which can be readily applied to data collected through surveys to obtain the above two results has been prepared by the Survey Research Center of the University of Michigan (Andrews *et al.*, 1967).

Extensive application has been made of this technique to data collected through sample surveys of 2,443 women, conducted in 1962 and 1963, just before and just after a nine-month program in Taichung City. An illustration is given in Table 9.10, which estimates the relative power of various combinations of demographic and social variables to explain the variation in family planning acceptance rates during the program period.

From Table 9.10 it can be seen that for the Taichung program the maximum percentage of the variance of the acceptance rate accounted for by all seven variables (three demographic and four social) is 16 per cent for all ages, 11 per cent for younger women and 22 per cent for older women. This indicates that even a set of crucial demographic and social variables leaves a large portion of unaccounted variance in family planning acceptance.

Table 9.10. *Proportion of variance in total acceptance rates explained by various combinations of demographic and social variables*

Characteristics	*Age of wife*		
	20-39	*20-29*	*30-39*
Three demographic variables ([a])	15	9	22
Three demographic variables and wife's education	16	11	22
Three demographic variables, wife's education and three other social variables ([b])	16	11	22
Three demographic variables plus wife's education and number of modern objects	16	11	22

Source: Freedman and Takeshita, 1969.

([a]) (1) Open birth interval and whether ever used family limitation before
(2) Number of living children and sons
(3) Comparison of number of children alive and wanted

([b]) (1) Husband's employment status
(2) Ethnic background
(3) Family type.

One of the essential conditions for the application of MCA is that there should be no interaction effects on the dependent variable. But as indicated earlier, interaction effects do exist in KAP/fertility variables. The procedure which is usually adoped to overcome this difficulty is to regroup the predictor variables into a new set in such a way that the interaction effects of the new set become insignificant. This is possible by first studying the interaction effects of all subgroups of the predictor variables and later rearranging them in such a way that new "interaction" variables are formed. These interaction variables are then used instead of the raw predictor variables. Sonquist and Morgan (1964) have developed a statistical procedure and a computer algorithm for automatic detection of interaction (AID) between independent variables. Using this program, it is possible

to find a set of predictors which maximize the reduction of error In the dependent variables. The AID program is sequential, generating a tree of two-way splits on the predictor variables starting with the entire sample. A dependent variable Y is chosen, and thenceforth everything turns on the search for the predictor that will maximize the between sum of squares (BSS) of Y on each occasion when the group is divided into two. Thus, AID scores every two-way split on every predictor, finds the best one, and so divides the sample into two smaller groups. The same procedure is continued for each of the new sub-groups and the process goes on till certain other criteria are satisfied. Application of the AID computer program to identify the best predictor variables (in a sequential order) on the rates of acceptance by women of family planning methods in a two-year period was made by Ross and Bang (1966) on the data collected in a follow-up study of a sample of women in Koyang and Kimpo in Korea. Among 21 variables included in the analysis, the open birth interval was found to be the best predictor for subsequent acceptance. Usually such an application of the AID program for study of interaction effects precedes the application of the MCA program for studying the effects of each predictor variable on the dependent variable.

IX. OPERATIONAL PROBLEMS AND NON-SAMPLING ERRORS

OPERATIONAL PROBLEMS

Any survey operation can be considered as an interrelated set of activities, technical and administrative, phased over time. Efficient performance of these activities in their proper time sequence is fundamental to the success of the survey. Preparation of a flow diagram indicating the various stages of the survey from planning to cleaning and analysis of data will greatly assist in streamlining the survey operation. An illustration of such a flow diagram prepared for the national "impact survey" of Pakistan is given in Figure 9.2 (Hardee and Sirageldin, 1969). The figure illustrates the phasing and interrelation of different stages from budget allocation to data cleaning as envisioned by the organizers at the beginning of the survey.

The operational problems can be considered at three broad phases: planning, field operations and data processing. In activities at each of these phases there is an overlap between technical and administrative considerations and a need for the survey director, survey specialist and statistician to work together in reaching decisions on significant problems.

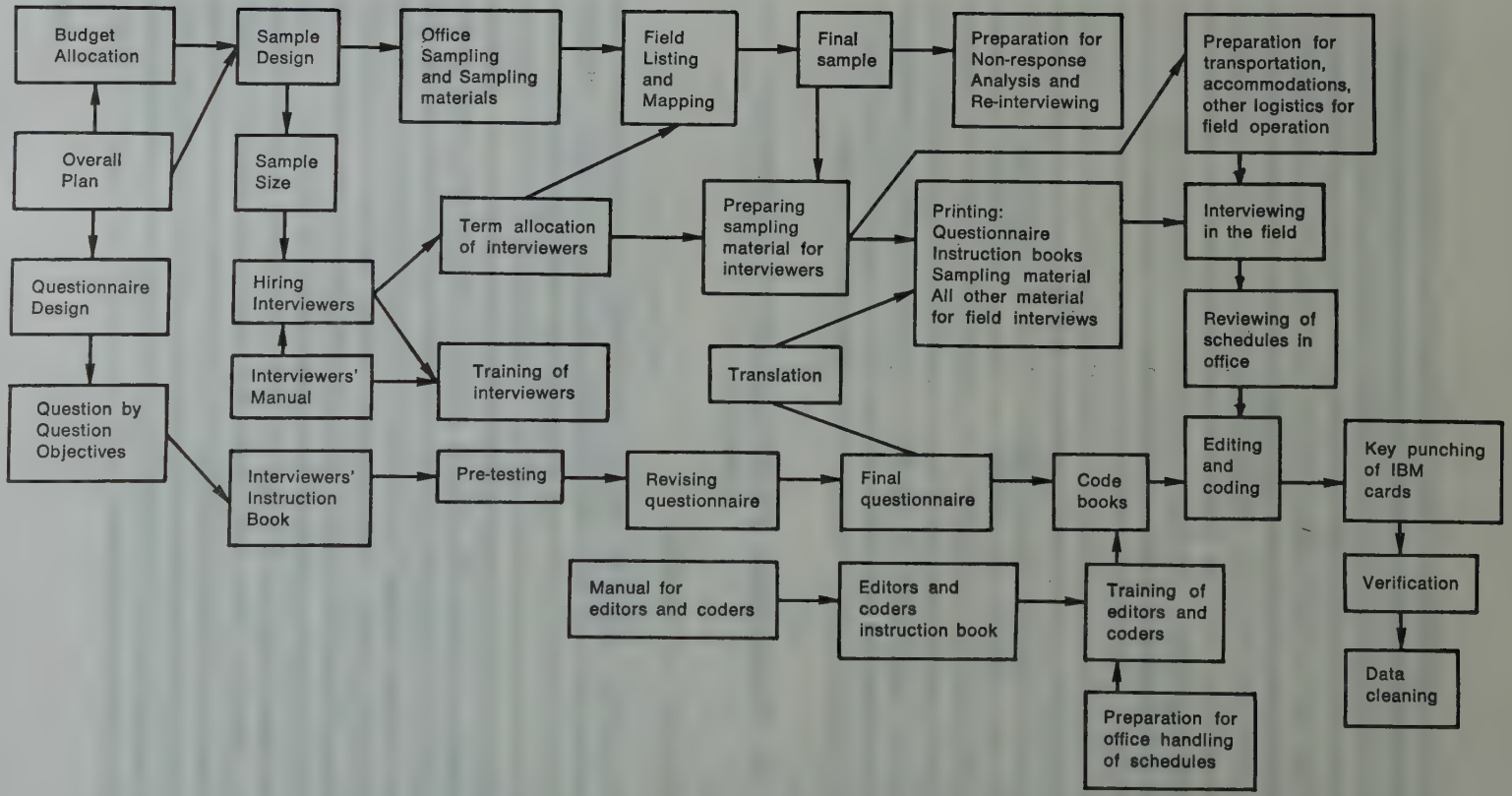


Fig. 9.2. *Flow diagram of field survey stages from planning through clean data (as designed from Pakistan's "impact" survey, 1968)*

NON-SAMPLING ERRORS AND BIASES

One of the serious criticisms levelled against KAP surveys is that the reliability and validity of data collected in these surveys are poor. Hauser (1967) argues that,

> "KAP surveys, along with many other sample surveys, have exhibited two major deficiencies. First, they have, in general, failed to include adequate efforts to study their reliability of data; and second, they have failed to make adequate efforts to obtain measurements of the 'intensity' of the opinions or attitudes reported . . . it is a moot question whether surveys do not mislead as much as they inform . . ."

Mauldin *et al.* (1970) observe that,

> ". . .KAP surveys score poorly in terms of reliability of individual responses, but rate highly in terms of consistency of aggregate distributions over time."

The errors and biases observed in KAP/fertility surveys have three components: due to respondent, due to interviewer and due to questionnaire design. Very few studies have attempted to estimate the magnitude of these errors, but the few that have been done reveal that the reliability of individual response on KAP items can be very poor. The Mysore Population Study (United Nations, 1961) attempted to measure interviewer differences in Household Survey data by a designed allocation of sample households to each interviewer in different strata, survey round and time interval, according to a Latin square design. The data collected were subjected to analysis of variance on crucial variables to study the significance of between-interviewer differences. The differences between interviewers on quantitative variables was not found to be significant. Systematic interviewer biases were observed, on the other hand, in the areas of opinions and attitudes toward family size and contraceptive practice.

There seems to be a tendency to under-report knowledge of contraception and contraceptive practice on the part of the respondents in developing countries. An investigation carried out in Bangladesh in 1968 on the basis of a follow-up of a sample of known users of contraception, selected from clinic records, revealed that "over one-third of the known users of contraceptives indicated they had never used, while less than 10 per cent indicated they had no knowledge of family planning" (Stoeckel and Chowdhury, 1969). It was also found that the extent of under-reporting varied inversely with the duration of use but was not related to the socio-economic status of the couples.

Thus, when one analyzes the changes in the proportion having knowledge or practicing contraception over time, the possibilities of response errors confounding real change should be taken into account in the interpretation of data. Built-in consistency checks on the response and reinterviews of a sample of respondents by better quality interviewers should form a regular part of the survey.

The reliability of response in relation to questionnaire design, especially the extent of probing adopted, has been observed in many surveys. For example, the data from the national impact survey in Pakistan revealed that the proportion of couples having knowledge about family planning methods increased consistently with the specificity of the question. This is brought out in Table 9.11.

With regard to knowledge of abortion, less than 1 per cent of the wives interviewed in the above study mentioned abortion when asked to mention all the methods of family planning (modern and traditional) they knew. Later, however, when the interviewer read a complete list of all methods, including abortion, and enquired about the respondents' knowledge, the proportion recognizing abortion as a family planning method increased to 38 per cent (Sirageldin, 1972). This indicates that the respondent did not consider abortion as a family planning method until it was mentioned by the interviewer.

Table 9.11. *Percentage of wives knowing about family planning and family planning methods in response to various questions for 5,998 women, 1969*

Type of question asked	Percentage who know about family planning				
		Bangladesh		Pakistan	
	All	Rural	Urban	Rural	Urban
1. General question about family planning	50	48	78	37	83
2. Specific question about family planning knowledge	62	51	71	75	84
3. More specific question about family planning methods:					
Knowledge about:					
Any method	78	64	82	97	98
Any modern method (*a*)	73	63	81	83	92
Any traditional method (*b*)	59	44	51	96	96

(*a*) Modern methods include condom, diaphragm, foam, jelly or cream, tampon or sponge, IUD (loop), pill, male and female sterilization and abortion.

(*b*) Traditional methods include abstinence, rhythm, douche, withdrawal and breast-feeding.

Source: Pakistan National Impact Survey (Hardee and Sirageldin, 1969).

The magnitude of response errors even in objective, measurable demographic variables such as age, number of children born, marriage duration, pregnancy status, etc., is found to be significantly high in developing countries. The contributing factors are ignorance, recall lapse and wilful misreporting on the part of the respondents. Studies on age misreporting because of digital preferences and other factors have been made by various authors. Comparisons of ages of persons as reported on successive occasions to two independent interviewers have revealed large variances increasing with age (Hashmi and Alan, 1969). The presence of such errors, even in carefully designed studies where special measures were taken for getting at accurate estimates of age, is quite depressing. The element of randomness in the reporting of such variables in surveys in developing countries should be constantly kept in mind in the analysis and interpretation of data.

In the reporting of live births and other types of termination of pregnancy, errors of omission of such events are found to be a function of the recall period; the longer the recall period, the higher the omission rate. Errors which are functions of the recall period have been termed as "recall lapse" and methods of estimation and correction for such "recall lapse" have been developed by Som (1972).

RANDOMIZED RESPONSE TECHNIQUE. During the last few years there has been increased use of the randomized response technique (RRT) to obtain reliable response on sensitive questions, such as the incidence of induced abortion. This technique was developed by Warner (1965) to assist in obtaining valid answers to questions that respondents may be reluctant to answer in an interview situation. The technique makes it possible for the respondent to provide truthful information on a sensitive or highly personal question without revealing his opinion to the interviewer. The sensitive question is mixed up with one or more insensitive questions for which the answer is known already, in known proportion, and the respondent is asked to choose one of these questions at random and give the answer "yes" or "no" without revealing the question which has been chosen. The technique has been used in North Carolina to detect abortion rates, oral contraceptive use and the proportion of emotional problems in a selected sample of women. It was also applied to one of the two sub-samples of currently married women in the ages of 15-49 in Taiwan (Chang, *et al.*, 1972). In the other subsample, data were obtained by the usual pregnancy history approach

to detect the proportion of women among the sample who had had at least one induced abortion. The estimated proportion according to the RRT technique was 28.2 per cent, compared with 12.7 per cent reporting induced abortion among those questioned in a direct and usual manner. The study also provides estimates of the proportion who did not cooperate with the RRT technique and it was found that this proportion was higher in the lower socio-economic group. This study reveals the potentiality of this technique for obtaining data on sensitive topics. However, since information for each member is not available through this technique, any analysis of association of frequency of abortion with individual characteristics of the woman or family would not be possible. Further methodological work in this field and field applications in different cultural backgrounds would be extremely valuable.

X. SUMMARY AND THE PRIORITY AREAS FOR FURTHER RESEARCH

In this section a summary is presented of the significant issues discussed in the previous sections on the role of sample surveys in the assessment of the demographic impact of family planning programs. Some of the priority areas for further research in this field are also indicated.

1. Sample surveys are important tools for the collection of the data necessary for demographic description, for understanding factors influencing demographic components and for family planning program monitoring and evaluation. They can provide data which are not obtainable from other more usual sources of demographic and family planning statistics, such as census, vital registration and family planning service statistics. There has been a remarkable increase in the number of surveys undertaken since 1960 for the study of human fertility and of family planning knowledge, attitudes and practices. It has been estimated that the number of such surveys is around 2,000, but most of these have been conducted in selected local areas, guided essentially by researchers from academic institutions and universities with the objective of obtaining a better understanding of sociological and psychological factors affecting fertility and contraception. In recent years the trend is shifting towards national or even world-wide surveys under the auspices of national governments and/or international organizations.

2. The number of fertility-related surveys of national or major geographical coverage undertaken since 1960 in developing countries

has been estimated at around 200. They can be classified into three types: multi-purpose, demographic, and fertility/KAP surveys, depending upon the objectives. The fertility/KAP type of survey has played a significant role in the context of national family planning programs, especially in gaining public support for such programs and to a limited extent for program monitoring and evaluation.

3. The objectives with which fertility/KAP surveys have been undertaken in the past have varied widely, ranging from pure scientific interest in understanding the sociological and psychological process underlying contraception and fertility behavior, to political motives for gaining support or funds for family planning programs. These surveys have revealed that even in the remotest rural areas of developing countries, women desire smaller families than they usually have or are likely to have, that they have favorable attitudes to contraception and family planning, and are willing to use modern methods of contraception if services are made available in a socially acceptable manner. Such findings came as a surprise to many social scientists who had believed that the prevailing high fertility levels in developing countries are entirely due to the desire for large families among couples in these areas.

4. Having helped establish the nature and magnitude of the population problem, the desire of couples to limit family size, and their willingness to practice family planning methods, KAP surveys are shifting their emphasis to monitoring and evaluation of family planning programs. In this direction the surveys are facing considerable difficulties both from methodological and operational points of view. Methodologically, the factors influencing adoption of contraceptive behavior by couples are not fully known. Even in a situation where all the demographic and social variables considered, *a priori*, relevant to contraceptive practice have been taken into account, they have been able to explain less than 25 per cent of the variance in contraceptive adoption. Similarly, the sociological and psychological variables known to affect fertility are able to explain less than 20 per cent of its variance. A large proportion of the variance in fertility and contraceptive behavior remains unexplained. Under these conditions, it is difficult to measure in precise terms the impact of family planning programs on levels of acceptance and fertility without recourse to actual experimentation using scientific principles of experimental design. Analysis of data collected through retrospective surveys of a sample population in an ongoing program can at best provide measures of association between different variables but cannot identify and assess the cause-effect relationship between program variables and acceptance of fertility.

5. However, using a conceptual model of factors influencing adoption of contraception and fertility behavior, it is possible to test a variety of hypotheses on the relationship between these variables using data collected from sample surveys. Many sample surveys have exploited this approach and a variety of conceptual models have been developed. They have usually been specific to particular disciplines of sociology, psychology and economics, and very few have attempted to synthesize all the variables into one holistic approach. If one appreciates the point that human behavior in vital areas such as reproduction and fertility is likely to be influenced by a host of factors overlapping the various disciplines of behavioral sciences, the need for such holistic causal schemes becomes readily apparent. Such an interdisciplinary integrated approach for the explanation of fertility and contraceptive behavior has been emerging in the last few years but further research is needed in this direction.

6. One of the fundamental gaps in the evaluation of family planning programs in developing countries is the lack of reliable data on fertility levels and trends. Sample surveys have assisted enormously in filling this gap. However, there appears to be a need for the development of sensitive indicators of fertility which will reflect short-run changes in period fertility measures. The extent to which random forces operate in the explanation of the observed variance in fertility also needs further study.

7. There appears to be a need for repeated surveys of the same population, either on independent samples or through panel studies, to study time trends in fertility/KAP variables and to gain a better understanding of the causal mechanisms behind these trends. The quality of data can be sequentially improved in panel studies.

8. It is disconcerting to note that different methods of statistical treatment of data collected from a survey might sometimes lead to different conclusions. Validity of the findings would be enhanced if the same results were achieved through different analytical methods. This points out the need for subjecting at least a part of the data collected to a variety of different techniques of analysis.

9. The problems of sampling and tests of statistical significance changes in KAP/fertility variables require further theoretical and empirical work. The sample sizes usually used in KAP/fertility surveys are not adequate for measuring, with reasonable statistical confidence, changes in fertility levels of less than 10 per cent in any given age group. They are, however, adequate for measuring changes in knowledge of and attitudes towards contraceptive behavior, which seem to change more rapidly than fertility.

10. The major criticisms levelled against fertility/KAP surveys are that the reliability and validity of the data collected in such surveys are poor. Systematic procedures for the improvement of the quality of data collected in surveys are of fundamental importance before any meaningful interpretation of the data collected can be made. The recent developments in methods of data collection, such as randomized response techniques, should be tested for their empirical validity in different situations.

Appendix 1. *Particulars of national fertility/KAP surve*

Number	*Country*	*Year*	*Sponsoring organization*	*Title*	*Type of s*
1	India	1970-1971	Central Statistics Organization	Special enquiry on birth and death reporting	Fertility
2	India	Continuing since 1962 6-monthly surveys in each area	Varies per survey	Standard fertility survey	Fertility
3	Indonesia	1968	Population Institute, University of Indonesia	KAP Study of Djakarta and environs	KAP
4	Iraq	1972	Central Statistics Organization	National Fertility Survey	Fertility
5	Jordan	December-February 1971-1972	Population Statistics Section/UNFPA	National Fertility Sample Survey	Fertility
6	Korea (South)	1968	Ministry of Health and Social Affairs (Seoul)/ Family Planning Federation of Korea	Fertility and Family Planning Survey	Fertility/KA
7	Korea (South)	1971	Korean Institute of Family Planning/Korea Population Institute	Fertility Abortion Survey	Fertility/KA
8	Korea (South)	1970-1971	Korean Institute for Research in the Behavioral Sciences/AID	Korean Attitudes and Birth Control Behavior	Fertility/KA
9	Lebanon	October 1972	Central Directorate of Statistics	Fertility Survey	Fertility
10	Malaysia	Planned 1970-1971	Department of Statistics	PES/KAP	PES/KAP
11	Pakistan	1968	Family Planning Council of Government of Pakistan	National Impact Survey	Fertility/KA
12	Syria	1972	Central Bureau of Statistics/UNFPA	National Fertility Survey	Fertility

(or planned) in countries of Asia during 1968-1974

e of oach	*Area coverage*	*Population coverage*	*Sample size*	*Comments*
nd ctive	Regional (5 states)	Total population	Population in 108 sample villages	25th round
nd ctive	5 or 6 rural areas near large urban center	Ever married females 15-45	2,000 households in *each* area; sample within areas	Reports are available
ound ctive	Djakarta and environs	Ever married men and women aged 15-49	2,241 males and females	Draft report and unpublished data available
ound ctive	National			Planned
ound ctive	National	Ever married women under 50	4,800 households	Data being analyzed
ound ctive	National	Currently married women under 50	5,400 women	Published book available
ound ctive	Regional (selected urban and rural areas)	NA	5,022 females	
ound ctive	National	Currently married women under 44 years of age	1,883 females and purposive sample for further testing	Book published
ound ctive	NA	Ever married women 15-44 years	6,000 women	Final report available
ound ctive	West Malaysia only	Total population	20,000 households	Followed 1970 Census preliminary KAP published
ound ctive	National	Total population with emphasis on females	6,000 ever married women and 3,000 husbands	Paper available, manual to be published shortly
	National	Total population	NA	Planned

Number	Country	Year	Sponsoring organization	Title	Type of su
13	Taiwan	January-February 1970	Taiwan Provincial Institute of Family Planning, Taichung	Nationwide Fertility Survey (Third Islandwide KAP Survey)	Fertility/KAP
14	Thailand	1968	National Research Council	Pilot Study of Family Health in Thai Muslim Communities in South Thailand	Fertility/KAP
15	Thailand	1969 and later	Chulalongkorn University Population Research and Training Center/ Population Council	National Longitudinal Survey of Social, Economic and Demographic Change in Thailand	Demographic KAP

ued)

approach	*Area coverage*	*Population coverage*	*Sample size*	*Comments*
dinal und ctive	Plain area of Taiwan	Currently married women 22-41 years of age	3,302 women	Two samples total 2,754 plus independent sample of 548 to check on KAP for main sample, preliminary article published
ound ctive	Regional (Vala province)	All married women and 30 per cent of their spouses	655 males and females	Report published
und ctive	National	Total population/ fertility items for currently married women	*Rural* = 1,500-2,000 households *Urban* = 2,000 households	Beginning in 1969 and alternate years thereafter (rural). Beginning in 1970 and continuing in alternate years (urban)

Appendix 2. *Content of model KAP questionnaires*

A. General Structure

	IUSSP	*Population Council*
Pages	32	22
Household questions	12	10
Numbered interview questions	114	48
Total questions (including parts)	233	132
Fertility (I) *	42	25
Family planning		
Knowledge, general (IV)	49	5
Knowledge of government program (IV)	—	4
Attitudes on family size (II)	14	10
Attitudes on family limitation, general (III)	19	14
Attitudes on family limitation, government program (III)	—	2
Attitudes on related variables (VII)	—	3
Practice (V)	26	10
Socio-economic characteristics (VI)	83	59

* Roman numerals in parentheses keyed to IUSSP list in *Variables for Comparative Fertility Studies.*

B. Detail of questions on practice of contraception, abortion, sterilization

	IUSSP	*Population Council*
Total questions	26	10
Contraception only		
Traditional	1	—
Modern	—	2
Sterilization only	3	—
Abortion only	9	5
Generally any of these		
Current Use	1	1
Current Method	3	1
Ever used	9	1

Source: Watson et *al.*, "KAP Surveys and Evaluation of Family Planning Programmes", paper presented for the ECAFE "Regional Seminar on Evaluation of Family Planning Programmes", held in Bangkok, Nov.-Dec. 1969.

REFERENCES

Andrews, G., J. Morgan and J. Sonquist

1967 "Multiple Classification Analysis", Institute for Social Research, University of Michigan.

Baum, S., C. Hay and J. Huguet

1972 "An Inventory of Fertility (and related) Surveys Conducted or Planned since 1968: Parts I and II", paper presented to the first *Ad-hoc* Technical Advisory Meeting of the World Fertility Survey, the Hague, 11-14 July.

Berelson, B.

1967 "KAP studies in fertility" in B. Berelson *et al.* (eds.), Family Planning and Population Programs, Chicago, University of Chicago Press.

Central Family Planning Institute, New Delhi

1967 "Summary Record of the Meeting of the Expert Group on Standard Fertility Surveys", (mimeographed).

Chang, C. I., L. P. Chow and R. V. Rider

1970 "The randomized response technique as used in the Taiwan outcome of pregnancy study", Studies in Family Planning, Vol. 3, No. 11, The Population Council.

Davis, K.

1967 "Population policy: Will current programmes succeed?", Science, No. 158, November.

Davis, K. and J. Blake

1956 "Social structure and fertility: an analytical framework", Economic Development and Cultural Change, 1956.

Duncan, W. G.

1973 "The nature and content of fertility surveys conducted throughout the world since 1960", Occasional paper of the World Fertility Survey, No. 1, October.

Easterlin, R. A.

1969 "Toward a socio-economic theory of fertility: . . ." in S. J. Behrman *et al.* (eds.), Fertility and Family Planning: A World View, University of Michigan Press.

1972 "The economics and sociology of fertility: A synthesis", paper presented to the seminar on Early Industrialization, Shifts in Fertility and Changes in Family Structure, Institute for Advanced Study, Princeton, 11 June-9 July.

Fawcett, J. T

1970 Psychology and Population. An occasional paper of the Population Council, New York.

Freedman, R.

1961-1962 "The sociology of human fertility: a trend report and bibliography", Current Sociology, 10-11 (2). Basil Blackwell, Johnson Reprint Corporation, third reprint, 1967.

Freedman, R. and J. Y. Takeshita
1969 Family Planning in Taiwan, Princeton University Press, Princeton.

Hardee, Gilbert and Ismail Sirageldin
1969 "A summary statement on Pakistan's national IMPACT survey", in N. Sadik *et al.* (eds.), Population Controls: Implications, Trends and Prospects, Proceedings of the Pakistan International Family Planning Conference, Dacca., Pakistan Family Planning Council, Islamabad.

Hashmi, S. S. and I. Alan
1969 "The problem of obtaining age data in Pakistan", paper presented at the International Union for the Scientific Study of Population Conference, London, September.

Hauser, P. M.
1967 "Family planning and population programs", a book review article, Demography, Vol. 4, No. 1.

Hermalin, A. I.
1968 "Taiwan: An area analysis of the effect of acceptances on fertility", Studies in Family Planning, No. 33, The Population Council.

Hill, Stycos and Back
1959 The Family and Population Control. University of North Carolina Press.

Jahoda, M., M. Deutsch and S. W. Cook
1951 Research Methods in Social Relations. Dryden Press, New York.

Kiser, C. V.
1962 "The Indianapolis study of social and psychological factors affecting fertility" in C. V. Kiser (ed.), Research in Family Planning, Princeton, Princeton University Press.

Kiser, C. V. and P. K. Whelpton
1953 "Resumé of Indianapolis study of social and psychological factors affecting fertility", Population Studies, No. 7.

Lewis-Faning, E.
1949 "Report on an inquiry into family limitation and its influence on human fertility during the past fifty years", paper of the Royal Commission on Population, Vol. 1.

Loewenthal, N. H. and A. S. David
1972 "Social and economic correlates of family fertility: An updated survey of the evidence", Research Triangle Institute, North Carolina, June.

Marks, E. S., W. Seltzer and K. Krotki
1972 Population Growth Estimation: A Hand-book of Vital Statistics Measurement. The Population Council, New York (mimeographed version).

Mauldin, P., W. B. Watson and L. F. Noe
1970 "KAP surveys and evaluation of family planning programs", unpublished paper, The Population Council, New York.

Nortman, D.
1972 "Population and Family Planning Programs: A Fact Book", Reports on Population and Family Planning, The Population Council, New York.

Perrin, E. B. and M. C. Sheps
1964 "Human reproduction: a stochastic process", Biometrics, Vol. 20.

Population Council
1970 "Manual for Surveys of Fertility and Family Planning Knowledge, Attitudes and Practice", New York.

Potter, R. G.
1970 "Births averted by contraception: an approach through renewal theory", Theoretical Population Biology, 1(3), November.

Ross, J. and S. Bang
1966 "The AID computer program, used to predict adoption of family planning in Koyang", Population Studies, Vol. XX, No. 1, July.

Schultz, T. Paul
1971 "An economic perspective on population growth" in Rapid Population Growth, National Academy of Sciences, Johns Hopkins Press.

Seltzer, W.
1970 "Measurement of accomplishment: The evaluation of family planning effects", Studies in Family Planning, No. 53, The Population Council, New York.

Sheps, M. C.
1967 "Uses of stochastic models in the evaluation of population policies", Fifth Berkeley Symposium on Mathematical Statistics and Probability, IV.

Sheps, M. C. and J. A. Menken
1973 Mathematical Models of Conception and Birth. University of Chicago Press.

Sirageldin, I.
1972 "The survey method in family planning research and evaluation: Review and a proposal", Paper presented to the Joint American Association for Public Opinion Research/World Association for Public Opinion Research Conference, Atlantic City, May.

Som, R. K.
1972 Recall Lapse in Demographic Enquiries. Asia Publishing House, Bombay.

Sonquist, J. A. and J. N. Morgan
1964 "The detection of interaction effects: A report on a computer program for the optimal combinations of explanatory variables", Monograph No. 35, Survey Research Center, University of Michigan.

Srikantan, K.
1967 "Effects of neighbourhood and individual factors on family planning in Taichung", Doctoral dissertation, Department of Sociology, University of Michigan, Ann Arbor.

Srinivasan, K.
1971 "Study of fertility change in developing countries through analysis of open birth intervals", paper presented to the Technical Meeting on Methods of Analyzing Fertility Data for Developing Countries", organized by the United Nations in Budapest, June.

Srinivasan, K., A. Muthiah and S. Krishnamoorthy

1969 "Analysis of the declining fertility in Athoor Block", Bulletin of the Institute of Rural Health and Family Planning, Candhigram, Vol. IV, No. 3, June.

Stoeckel, J. and M. A. Choudhury

1969 "Pakistan: Response validity in a KAP survey", Studies in Family Planning, No. 47, November, The Population Council.

United Nations

1961 The Mysore Population Study. Population Studies, No. 34, (ST/SOA/Series A/34), New York.

1970 Variables and Questionnaire for Comparative Fertility Surveys. Population Studies No. 45 (ST/SOA/Series A/45), New York.

Venkatacharya, K.

1969 "Certain implications of short marital durations in the analysis of live birth intervals", Sankhyā, 31B, pp. 53-68.

Warner, S. L.

1965 "Randomized response—a survey technique for eliminating evasive answer bias", Journal of the American Statistical Association, 60: 63-69.

Watson, W. B., W. P. Mauldin and L. Z. Frank

1969 "KAP surveys and evaluation of family planning programs", paper presented to the ECAFE Regional Seminar on the Evaluation of Family Planning Programmes, Bangkok, November-December.

West Malaysian Family Survey

1970 West Malaysian Family Survey, 1966-67: Report by the National Family Planning Board, Kuala Lumpur.

Chapter 10
SIMULATION STUDIES

by
Jane Menken

I. INTRODUCTION

Demographers have long developed models which imitate or simulate the processes they are studying. The oldest example is the life table which both describes and simulates the mortality experience of a cohort. A variety of purposes are served by these models. They may imitate reality in that real data can be "fitted" or summarized by a specific model. More important, once a model is developed, it provides a vehicle for carrying out pseudo-experiments. For example, changes in life expectancy at birth can be studied by altering the patterns and levels of age-specific mortality rates in the life table model. More complex demographic models have been developed in response to the impossibility of performing adequate laboratory or controlled experiments in human populations and the consequent need to develop alternate procedures for studying interrelationships among factors which are believed to affect demographic processes.

This chaper ([1]) will be concerned with models in which the quantitative effects on reproduction in a population of such factors as mortality, marriage patterns, biological determinants of fertility, use and effectiveness of contraception or other methods of fertility control, and family size desires can be investigated. These models share the feature that the variables included are all treated as characteristics of an individual or couple and as direct determinants of their reproductive behavior. Since the effects of socio-economic and psychological factors on fertility operate through intermediate biological variables (Davis and Blake, 1956), these factors are not explicitly

([1]) The chapter is an extension of Mindel Sheps' paper "A review of models for population change" (1971).

included in most of the models; nor do characteristics of the administrative organization of family planning programs, such as the number of health workers employed, enter explicitly; rather, they may be included implicitly through changes in the probabilities of accepting a contraceptive method. The models described here do not, therefore, provide direct estimates of cost or cost-benefits of a program.

A simulation model "specifies a hypothetical population or subgroup and incorporates rules or assumptions for bringing the population forward in time through the occurrence of demographic events such as conceptions, births, deaths or marriages" (Sheps, 1971). Models of this type have both substantive and methodological applications to problems of family planning policy decisions and evaluation. Specifically, models can serve:

TO FOCUS ATTENTION ON GAPS IN SUBSTANTIVE KNOWLEDGE

Construction of a model requires precise conceptualization and detailed hypotheses concerning the variables to be included and the interrelationships among them. This, in and of itself, can focus attention on deficiencies in our knowledge of the process under study.

TO ESTIMATE THE RELATIVE IMPORTANCE OF FERTILITY DETERMINANTS

In the artificial environment shaped by the structure of the model, the relative importance of the factors which affect birth rates may be studied. These studies in turn may lead to suggestions for basic studies in real populations of factors which are judged crucial. They may also suggest types of family planning strategies which could be applied in real situations.

TO ESTIMATE THE RANGE OF POSSIBLE VALUES OF A SPECIFIC FERTILITY DETERMINANT

If all parameters but one in a fertility model are known, the range of values for this last parameter can be estimated by making a series of calculations which first set the value of the parameter at what the researcher believes to be an unrealistically low minimum and then increasing it gradually to a maximum suspected to be larger than that occurring in any real population (Leridon and Jacquard, 1973). Values that lead to simulated birth rates or other results outside a range considered reasonable in a human population are rejected, thus setting the desired bounds.

TO SIMULATE ALTERNATIVE FAMILY PLANNING PROGRAMS

A proposed family planning program can be simulated to give planners some idea of the results to be expected if real life resembles the abstract model. A numerical example frequently has valuable heuristic purposes, offering a precise estimate of the implications of verbal proposals and often revealing unsuspected consequences. Also alternative family planning strategies can be examined to help decide which approach might prove more effective and to determine whether feasible targets in terms of acceptors and the duration of their contraceptive use can possibly produce the desired changes in overall numbers and rates of births.

TO EVALUATE AND COMPARE INDICES AND ANALYTICAL TECHNIQUES

Models have played an important role in the development of methods of measuring fertility and of evaluating family planning programs. Investigators studying actual populations are usually limited in their choice of measures, both by the availability of data and by the size of the group under study. In a simulation, data can be generated under known conditions from a sample limited in size only by the capacity of the computer. A suggested method of evaluation can be applied to these hypothetical data and the results compared with the known population parameters. Ideally, an evaluation procedure should lead to conclusions which will reflect the true population values and discriminate well between different sets of parameters (Ridley and Sheps, 1966).

TO STUDY SAMPLING VARIATION

Certain models, usually employing a technique called microsimulation, can provide some idea of the sampling variation to be expected in the indices used to evaluate family planning programs.

TO DELINEATE SHORT- AND LONG-TERM EFFECTS OF CHANGES IN REPRODUCTIVE BEHAVIOR

When new patterns of fertility control are accepted by a population group, the birth rate does not immediately drop to a new stable level; rather, there is a time-lag before any effect can be observed and many irregularities occur in the approach to a new relatively constant level. Simulation models can provide quantitative examples of the kinds of patterns to be expected, perhaps preventing either

overly optimistic or overly pessimistic reactions to the early results of a program.

TO ESTIMATE THE EVENTUAL IMPACT OF A CURRENTLY OPERATING FAMILY PLANNING PROGRAM

Finally, when a family planning program is at various stages of operation, preliminary data, such as acceptance and continuation rates, may be converted into estimates of the long-term effect of the current program. It must be emphasized that these results are only estimates and are accurate only insofar as the relationships between variables included in the model reflect reality and the estimates of the levels of independent variables and their trends are accurate. Simulation models have yet to provide very good "fits" to actual population data (Barrett, 1971; Holmberg, 1972) and must be viewed as approximations, perhaps crude ones, of demographic processes in these populations. Also, one need only recall the pre-baby boom predictions of population decline in western Europe and the United States to become convinced of the great caution necessary in treating results from population models as forecasts of the future. Flagrant misuse of models in the recent past has occurred through lack of awareness or deliberate ignoring of these facts. However, information from models may give a quantitative picture of the future consequences of the complex sets of assumptions or observations derived from a family planning program and may, therefore, aid in reaching decisions about the continuation or alteration of the program.

To facilitate discussion of models and their applications, section II of this chapter consists of a description of a simple model and a classification system based on characteristics of models. Examples of applications to family planning programs appear in section III, followed by a discussion in section IV. Because the literature on demographic simulation models is scattered and frequently difficult to locate, a brief review of existing models appears as an Appendix.

II. SIMULATION MODELS

The models already developed or proposed vary considerably in their structure, partly due to differences in their purpose and partly because of variations in techniques of approach. As a basis for description and comparison of the different types of models, the initial illustration will be a very simple one.

AN EXAMPLE

BIOSIM (Sheps and Menken, 1973, Chapter 6) incorporates a minimum of variables under quite restrictive assumptions. The model was developed in response to a particular problem: When a population with a relatively constant birth rate changes its reproductive patterns, these changes are reflected in the birth rate only after a time lag. If the new levels of such factors as contraceptive use are maintained over a long period of time, the birth rate will approach a new constant value. We were interested in studying the pattern of approach to this new level, especially as it related to the numbers of women in the population susceptible to conception. Only three major biological determinants of birth rates were considered: the probability of conception per month (fecundability), the pregnancy outcome, and the duration of the nonsusceptible periods that follow a conception. All susceptible women are assumed to have the same probability, $\rho(t)$ say, of conceiving in month t, but $\rho(t)$ can change with time. Therefore, changes in fecundability due to aging or to acceptance of contraception can be simulated. If a conception occurs in month t, the pregnancy may end in an abortion (spontaneous or induced) k' months later, i.e., in month $t+k'$, with probability $\alpha(t + k')$. The probabilities of fetal loss can be varied in the model from month to month; hence, changes in birth rates produced by increasing fetal loss due to aging or to use of induced abortion can be studied. If a fetal loss occurs in month $t + k'$, the woman can next conceive again in month $t + w$. In the case of a live birth, pregnancy ends at $t + k$, where k is usually taken as 9 months and the next conception may occur in month $t + m$ [2]. The computer program written to calculate numerical results for this model records for each month t the number of women who are of parity n (i.e., have had n pregnancies which ended in live births), $n = 0, 1, 2, \ldots$, and subdivides them into those who are

i) susceptible to conception,
ii) in each of the early months of gestation, i.e., in months $1, 2, \ldots, k' - 1$, before the pregnancy outcome is determined,
iii) in each month of the postpartum nonsusceptible period associated with a fetal loss,
iv) in each month of a pregnancy which will end in a live birth, and
v) in each month of the postpartum nonsusceptible period following a live birth.

[2] In this model, $m \geqslant w$, $w \geqslant k'+2$, $m \geqslant k+2$ and $k \geqslant k'$.

A diagrammatic representation of BIOSIM is shown in Figure 10.1.

The initial population is a marriage cohort of women which can be observed in one of two possible ways. Under the first option, we assume that all women are just entering a sexual union and are

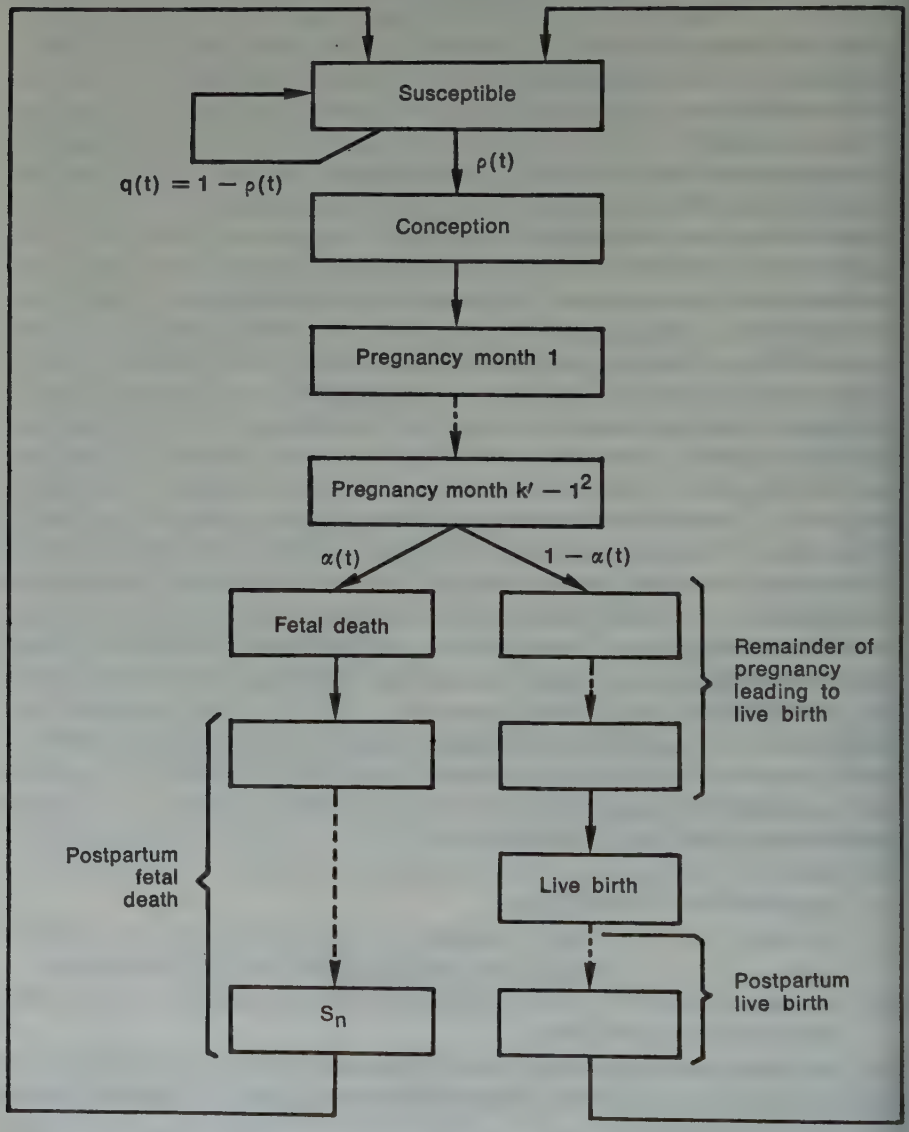


Fig. 10.1. *BIOSIM: A model with two pregnancy outcomes in which fetal death may occur k′ months after conception and all nonsusceptible periods have constant duration*

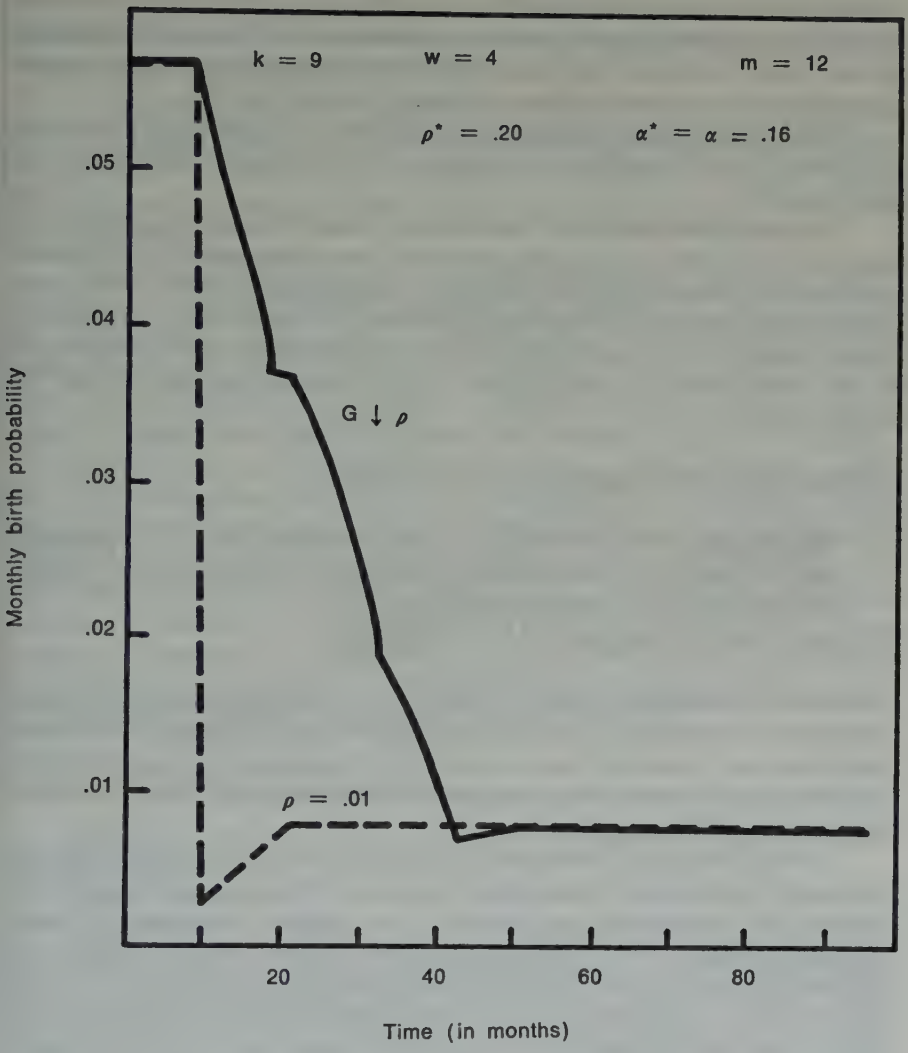


Fig. 10.2. *Birth probabilities starting from a constant value (original parameters are $\rho^*=0.20$ and $\alpha^*=0.16$) and assuming an abrupt or a gradual decline in fecundability*

susceptible to conception. In the second, the cohort is assumed to have been subject, over a long period of time, to constant values of the reproductive parameters described above. This leads to a situation where, at any point in time, the birth rate is constant, as is the

distribution of women among the states of susceptibility and the months of pregnancy, etc., previously discussed. Observation begins, in this case, a number of years after the formation of stable sexual unions and the cohort is then followed forward in time.

The output of the program includes the probability of a birth to a woman in the computer population and the parity distribution at each month.

In one experiment with this model, under the second option, it was assumed that over a long period of time fecundability had been constant and relatively high (0.20) in a population with a short non-susceptible period (3 months) following a live birth. Pregnancy lasted nine months, so that the total nonsusceptible period following a live birth was $m = 12$ months. Other parameters are shown in Figure 10.2.

Contraception was introduced in two ways:

i) fecundability abruptly fell to a very low level of 0.01, representing 100 per cent acceptance of a 95 per cent effective contraceptive with no discontinuation, or
ii) fecundability gradually fell to this level.

The second assumption can be thought of as the mathematical representation of a cohort in which the proportion of acceptors increases over time, or in which the proportion of acceptors is constant but the efficiency of their contraceptive use increases with time, or some combination of these patterns.

In the first case, the birth probability dropped precipitously in the tenth month after introduction because of the sudden drop in the number of conceptions at month 1. It climbed, however, in subsequent months, until it reached a new constant value. Why? At the start of the program, as a result of their high fecundability, many women were pregnant or in a postpartum nonsusceptible state so that the number susceptible to conception was low. With extremely effective contraception, few of these women conceived, so that more than usual remained susceptible until the second month and a number of women who had conceived m or w months earlier returned to susceptibility. Thus, at month 2, more women were susceptible than at month 1. Therefore, the number of births nine months later, at month 11, was greater than at month 10. For the next few months, the number of susceptible women (and consequently the number of births and the birth probability nine months later) continued to increase until a new steady state was reached. In the second case,

labeled $G \downarrow \rho$ on Figure 10.2, even though the fecundability was changing gradually for 33 months, the birth probability still dipped below its ultimate steady state value.

In other experiments, increasing the rate of fetal loss produced similar results. Also, when the reproductive parameters were altered to produce a higher birth probability, the monthly rate rose above the new level it would ultimately achieve and then gradually approached a constant.

Thus, even this extremely unrealistic model provided insight valuable to a family planning program: the results of a program may be misleadingly encouraging in the first few months unless strict attention is paid to the problem of identifying the number of women who are susceptible to conception at each month of the program. Potter, Masnick and Gendell (Potter and Masnick, 1971; Potter *et al.*, 1973) have explored this problem extensively. More generally, work with BIOSIM and similar models demonstrates that changes in reproductive behavior are not necessarily followed by a new constant birth rate after a lag. Rather some fluctuations are to be expected. As previously emphasized (Henry, 1972, p. 61), differentials and trends in birth rates should not necessarily be taken as indications of corresponding changes in fecundability or abortion rates. Inferences from levels of birth rates to levels of these underlying factors must be made with caution (Sheps and Menken, 1973).

CHARACTERISTICS OF MODELS

Simulation models may be classified according to the following characteristics:

1. *APPROACH TO FERTILITY.* The *demographic* approach treats the occurrence of a birth as a probability which may be specific for selected variables including age, marital status, parity, interval since the previous birth, etc. The *biological* approach explicitly includes the determinants of the birth interval, i.e. fecundability, probabilities of each type of pregnancy outcome, the durations of the period following conception when a woman is not susceptible to another conception, and, in some cases, parameters related to family planning. The former is useful primarily for tracing the consequences of fertility change for population change, the latter for assessing the results of changes in reproductive parameters (including the effects of family planning practice) on fertility.

2. *TYPE OF POPULATION.* The hypothetical population or group in a simulation may be (a) a *cohort* or a *cross-section* of a population, and (b) a *one-* or a *two-sex* population, and (c) *homogeneous* (all individuals in the group are exposed to the same probabilities of various events) or *heterogeneous* (probabilities may vary among individuals according to some distribution).

3. *THE MATHEMATICAL APPROACH TAKEN*, which may be *analytical* or *numerical.* Analytical models express assumptions about the individual variables in mathematical form and yield expressions summarizing the interrelationships among the factors and their effect on the dependent variable of interest. Solutions for these models have usually been obtained only under severe restrictions on the assumptions and with probabilities unchanged over time, as, for example, in stable population models. Analytical models may be either *stochastic* or *deterministic.* In the former case, if a population of size N is exposed to the risk, p, of a given event occurring, then the number of events which actually occur is a random variable with expected value Np (and variance Npq, etc.). By contrast, in a deterministic model, p is viewed as the proportion of the population to whom the event will occur, and therefore Np events are recorded. Deterministic models are frequently referred to as expected value models.

In a numerical model, either *macrosimulation* or *microsimulation* may be attempted. Because of the extent and complexity of calculations, numerical simulation models are almost always computer models.

a) A macrosimulation model involves mathematically formulated population structures. A population is divided into subgroups according to specified criteria and postulated probabilities of the occurrence of various events (in a stochastic macrosimulation model) or proportions (in a deterministic one) are allowed to operate on the subgroups for an appropriate period of time. Component projection of population is the best known example of a demographic macrosimulation technique and, like almost all models of this type, it is deterministic because proportions, not probabilities, are applied to determine the numbers of survivors and births in each category.

Macrosimulation models allow more realistic assumptions regarding the underlying biological factors affecting reproduction than has so far been possible with analytical models. Their development is hampered because the capacity of even the largest computers limits the number of categories that can be treated and the range and

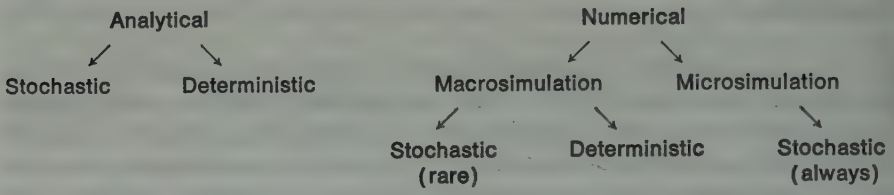
number of analyses or results that are possible. For example, if one wished to extend BIOSIM to include contraceptive acceptance dependent upon parity, then the program would have to subdivide women of parity n into two groups, acceptors and non-acceptors. For each group, all of the information on current status (e.g. susceptibility, month of pregnancy, etc.) would have to be recorded, thus, at a minimum, doubling the computer storage required and decreasing the computer space that can be allotted to program instructions. Introduction of other factors such as infant mortality or family size desires would mean further geometric increases in information to be stored by the computer, soon exceeding computer capacity or programmer ingenuity.

b) In contrast, microsimulation, which is always stochastic, builds up the population structure starting from primary demographic components, e.g. from individuals. A history for each individual in the model population is developed, in which at least some characteristics are treated as being determined by a random device. For example, consider a woman whose history has been developed up to the point where she is pregnant. A random number (R) between zero and one is generated by a special routine, and is compared with the predetermined probability (α) that pregnancy will end in fetal loss. If $R \leqslant \alpha$, pregnancy wastage is assumed to occur, and if $R > \alpha$, the pregnancy ends in a live birth. In contrast, BIOSIM, in addition to being a macrosimulation model, is deterministic. If N pregnancies occur at time t, $F = [\alpha(t + k')]N$ will end in fetal loss, whereas in a microsimulation model, the actual number of such pregnancy terminations is a random variable with expected value equal to F.

Many more variables can enter into this type of model, since only one person's history need be stored at a time. This history may be recorded on a tape or other form of output and all histories analyzed by a subsequent program. Distributions of age at marriage and sterility may easily be introduced. The physiological functions and their modifications by contraception, induced abortion, or surgical sterilization may be made dependent on age, parity, desired number of children, child mortality, etc. It is also relatively easy to provide for considerable interaction amongst variables. The analyses may be as detailed and varied as desired, since the output resembles a complete set of data from an ideal survey, and need not be limited to a restricted storage space. On the other hand, detailed data for input are not available for many of the most interesting variables and relationships so that the investigator finds himself in the frustrating position of making "educated guesses" for input values. Also, because

of the method of operating through random numbers, the results obtained constitute a random sample of possible results from repeated independent simulations. Hence, although this feature is an advantage in studying sampling variations when all parameters are held constant, in some cases large samples may be required in order to obtain reliable results. This consideration does not arise in deterministic models which calculate accurate results under the specified assumptions.

To summarize, the mathematical approaches to models may be classified in the following way:



4. *VARIATION WITH TIME.* A model in which probabilities are unchanged from cohort to cohort is *stationary in time.* None of the probabilities is a function of calendar or "external" time; however, in certain models, probabilities may vary with age, marital duration, or any other "internal" time factors. Such models tend toward a statistical equilibrium under suitable conditions. By contrast, a *dynamic* model permits secular change in probabilities and is, therefore, usually more appropriate for the detailed study of short-term effects of family planning programs.

The distinction between stationary and dynamic models is important only when the population is cross-sectional. Most cohort models that allow variation with internal time factors may be treated as either stationary or dynamic. For example, in BIOSIM, a sudden increase in the fetal loss rate beginning at time t can represent either a pattern of variation in the rate by marital duration common to all cohorts or the effect of a liberalized abortion law promulgated when this cohort had reached marital duration t. The effect of the law on the cohort with marital duration $t - 1$ could be studied in another simulation in which the fetal loss rate increased at $t - 1$. (If only cohorts of fairly long marital duration are affected by the change, the effect may be the same in terms of decline in cohort birth rates.)

5. *PROVISIONS FOR FAMILY PLANNING.* Family planning may be introduced *implicitly* through trends and changes in the reproductive parameters or in birth rates. In models directed more specifically

to this topic, an *explicit* structure for family planning is present. Acceptance of family planning may depend upon age, parity, desired family size, birth interval, contraceptive method, etc. Demographic models with explicit family planning structures introduce time trends in birth probabilities which may be specific for the variables just listed. In biological models, the changes affect the physiological parameters. Duration and effectiveness of contraceptive use, switching patterns and other characteristics of the family planning method can be introduced more easily into this type of model.

According to this classification system, BIOSIM is a deterministic macrosimulation model with a biological approach to fertility. Its population is a one-sex homogeneous cohort. Provisions for family planning are implied in the assumptions governing the reproductive patterns. If the cohort is treated as a sample from all cohorts having the same reproductive traits, the model is stationary. Otherwise, if $\rho(t)$ and $\alpha(t)$ are considered dependent upon calendar time, as in the examples given, the model is dynamic.

III. APPLICATIONS OF SIMULATION MODELS

The literature on demographic applications of simulation models is becoming voluminous. A wide range of studies employing simulation models in analyzing demographic interrelationships (not necessarily directly related to family planning), can be found in the references (cf. Henry, 1972; Perrin and Sheps, 1964; Sheps, Ridley and Lingner, 1966; Bodmer and Jacquard, 1968; Holmberg, 1970, 1972; Potter, 1972; Sheps and Menken, 1973). The examples in this section, selected from the many investigations especially pertinent to family planning, are grouped according to the types of applications suggested in section I, although not all are represented here. They are also intended to illustrate the sorts of questions which can be treated by different kinds of models.

ESTIMATES OF THE RELATIVE IMPORTANCE OF DIFFERENT FERTILITY DETERMINANTS

Sheps and Perrin (1963) reported an analytical model which related fecundability, contraceptive effectiveness, probability of fetal loss and durations of nonsusceptible periods to birth intervals and monthly birth rates. The model is deterministic, stationary, and homogeneous, i.e., all parameters are the same for all women, and the values do not change with age. Under these assumptions, if fecundability is high (.20), the probability of a fetal loss is 0.30, the

duration of pregnancy and postpartum nonsusceptibility following a fetal loss is 5 months, and a live birth occurs after nine months gestation, the birth rate is highly dependent upon the length of the nonsusceptible period following a live birth. The monthly birth rate to fecund women is given by

$$b = \frac{\rho(1-\alpha)}{(1-\alpha)\,\rho m + q + \alpha\,\rho w}$$

where

ρ = fecundability,
$q = 1 - \rho$,
α = probability that a conception will end in a fetal loss,
m = total nonsusceptible period (pregnancy + postpartum) associated with a live birth, and
w = total nonsusceptible period associated with a fetal loss.

If, in one population, the anovulatory period lasts three months, the monthly birth rate (to fecund women) a long time after marriage is 0.050 while in a second population with nonsusceptibility of one year, the rate is 0.035. In other words, in two highly fecund groups, the birth rate in the one with a long nonsusceptible period is 30 per cent below that of the population with a short period.

This model also quantified the differences in the long-term effects of contraception on these two populations. If fecundability were reduced by 95 per cent, to 0.01, the birth rate would be reduced by 87.2 per cent in the population with the short nonsusceptible period and by 82.5 per cent in the second population. Unfortunately, little is known about the true duration of the nonsusceptible period. The model might also be applied to estimate a range of values for this duration, as discussed in the Introduction.

This version of the Perrin-Sheps (1964) model of human fertility probably bears little resemblance to the reproductive process in a population where a family planning program is introduced, particularly because significant discontinuation rates of use among acceptors are recorded. Potter and his associates (Potter and Masnick, 1971; Potter *et al.*, 1973) have treated this problem in relation to the timing of acceptance and its overlap with postpartum nonsusceptibility in an analytical model which again illustrates the importance of determining susceptibility or the duration of anovulation in acceptors.

Family planning programs are intended to enable individuals to determine and achieve their chosen family size and spacing. Potter *et al.* (1968) asked whether or not couples practicing effective contraception risked the possibility of being unable to attain their desired

family size within a reasonable number of years after marriage. A model in which fecundability depended upon marital duration and parity was required. Analytical results are difficult to obtain in this situation. The investigation was pursued using FERMOD (Potter and Sakoda, 1966), a large-scale deterministic macrosimulation model representing a homogeneous marriage cohort of women with a reproductive period of specified duration and constant probabilities of miscarriage and stillbirth, as well as constant durations of gestation for each pregnancy outcome. Postpartum anovulation was given a constant duration following a miscarriage or a stillbirth, but a probability distribution following a live birth. Fecundability could be varied according to parity, thus permitting increasing contraceptive efficiency as desired family size was approached and attained. It could also be varied according to duration of marriage, allowing simulation of different spacing preferences. Thus, the parameters varied with internal time, so that this model could be considered either stationary or dynamic according to the researcher's purpose.

The expected number of conceptions, births, etc. occurred each month, but, as time went on, the hypothetical women were stratified in terms of their achieved parity, the time of their most recent birth and, if pregnant, the time of their most recent conception. Consequently, expected distributions of women by parity, and of intervals from marriage to the first, second, third, etc. birth could be obtained.

In one part of their study, these investigators postulated that couples wanted three children spaced at two, five and eight years after marriage and that they would use a 90 per cent effective contraceptive for spacing purposes, starting at marriage, and a 99 per cent effective contraceptive after the third child was born. If, after two years, no birth had occurred, contraception would be dropped until after the birth of the first child. Similarly, contraception was dropped if the second or third child was "late" in arriving. Postpartum periods were short (after a live birth, the mean duration was 3.5 months with a standard deviation of 1.6 months), and 20 per cent of pregnancies ended in either abortion or stillbirth. Under these conditions, when natural fecundability was 0.26, 96.5 per cent of couples had at least three children within 10 years of marriage and all had at least three within 20 years. Therefore, results from this model suggest that, among couples with relatively high fecundability, extensive contraceptive use would not prevent their having the desired three children.

In reality, however, fecundability varies among women. Without altering any other parameters, further experiments were carried out in which fecundability varied among women but the population mean fecundability stayed at 0.26. They demonstrated that the probability

of having at least three children within 10 years of marriage remained high (0.81) even when a high proportion of couples had natural fecundability less than 0.10. Thus, variation of this type does not markedly alter the results obtained under the simplifying assumption of homogeneous fecundability.

SIMULATION OF ALTERNATIVE FAMILY PLANNING PROGRAMS

Analytical and macrosimulation models have proved helpful in studying the demographic consequences of a family planning program to a cohort of acceptors. The capacity of computers has limited the numbers of variables and changes over time which can be included and has restricted the study of patterns of contraceptive acceptance and switching. Microsimulation models, flexible enough to include many variables with few restrictions, are more suitable for this purpose.

Sarma (1970) simulated alternative family planning strategies on POPSIM, a dynamic model designed to simulate all the principal demographic processes except migration (Horvitz *et al.*, 1971). Its population may be defined either as a one- or two-sex cohort or as a two-sex cross-sectional population. Rather detailed input data are required. The approach to reproduction is demographic. Birth probabilities vary with age, marital status and parity. Marriage is a function of age and current status (single, widowed, divorced). Mortality probabilities depend upon age, sex and marital status. Several methods of family planning can be simulated at once. The proportion of women, by age and parity, practicing a given method at the start of the program must be specified. The duration of use is determined by drop-out rates which are functions of age, parity and method. Probabilities of acceptance of a method and, when that method is dropped, of switching to each of the other methods or to no method are necessary. Use-effectiveness and the proportion of women using a method determine the reduction in the birth rate. The model is intended as a multi-purpose tool for studying a wide variety of demographic situations.

Sarma considered three components of family planning programs: eligibility, acceptance, and switching patterns and tested two levels of each of these factors. Eligibility depended upon numbers of living children, their age and sex; adoption of family planning was a function of family size and age of wife; patterns of switching to another method or of discontinuing altogether were specified for three methods (conventionals, IUD, and sterilization). All other parameters were estimates of demographic characteristics of India in 1960.

Ten years of program activity were simulated and five indices of fertility calculated. Without giving details of the programs tested, let us consider the kinds of output Sarma was able to examine. From the fertility indices, the short- and long-term effects of the program on demographic factors could be studied. By replicating a given program six times, Sarma was able to estimate the sampling error for these indices in a sample of a given size, a result impossible to obtain in a macrosimulation. The number of acceptors in each year of the program was also presented, providing an estimate of the workload for the family planning program and thus an estimate of the target number of acceptors for each year of the program necessary for the achievement of the calculated reductions in fertility. These figures also serve to predict needs for contraceptive supplies at successive stages of the program. Data pertinent to determining costs and personnel requirements can be taken from the simulation output. Although great caution must be exercised in their application, calculations such as these do provide the most complete analysis currently possible of the future consequences of a program without experimenting with real populations.

PROBLEMS OF DATA ANALYSIS

Simulation models provide an enormously powerful tool for studying methods of measuring levels and changes in fertility determinants. The hypothetical population is constructed from known variables and relationships and a method of analysis can be applied to the simulated data. A method which adequately reflects the values of the underlying variables and differentiates among different sets of parameters can then be applied with greater confidence to real data.

In recent years the open interval since the most recent live birth has been suggested as an indicator sensitive to new innovations in reproductive behavior. Sheps and Menken (1973, Chapter 9) demonstrated, via an analytical model, that the observed distribution of open intervals (and other variables measuring the duration between two events) depends upon the sampling frame (the selection process by which women enter into a sample) and may not adequately reflect underlying reproductive parameters. Open intervals were also simulated under different sets of assumptions on REPSIM-A (Ridley and Sheps, 1966), a microsimulation model with a biological approach to fertility, which will be described below. The ability of data on open intervals to distinguish among the sets of parameters was seriously questioned (Sheps *et al.,* 1970). Similar, more detailed,

conclusions were reached by Venkatacharya (1972) in an application of another microsimulation model.

Lee and Isbister, Wolfers, and Potter estimated births averted by a family planning program ([3]) by different methods. Clague and Ridley (1973) simulated two cohorts on REPSIM-B, a revised, more efficient version of REPSIM-A. REPSIM-B permits considerable variation in biological and demographic factors, including certain family planning assumptions. It deals with an age cohort of women, allows fecundity parameters to vary with age and between women, and also includes variation in age at entry into sexual union (marriage), at dissolution of marriage, at sterility, or at death. Provision is made for mortality of women, their husbands and their infants and for remarriage. Three types of fertility control are included; induced abortion, sterilization and family planning (defined as contraception, rhythm or coitus interruptus). Acceptance may depend upon parity or the gap between current parity and desired family size.

For the first cohort, with no fertility control, numbers of births and age-specific birth rates were calculated. The second cohort was exposed to a family planning program. An estimate of the "true" number of births averted in this hypothetical population may be obtained by differencing the numbers of births in the two simulations. The three different estimation procedures suggested by Lee and Isbister, Wolfers and Potter were applied in turn to data for the second simulated cohort. The resulting estimates of births averted were compared with the "true" value and with each other. Because they simulated only one population under one set of assumptions, the authors have yet to conclude which of these methods is the best but present their results as an illustration of the uses of simulation models.

EVALUATION OF PLANNING PROGRAMS

Chandrasekaran *et al.* (1971) criticize current usage of the concept of births averted because it lacks a time dimension. Estimates are for a segment of use or for the cohort's remaining reproductive period. Their criticism is that current procedures fail to allow for postponement of births, i.e., for the "making up" in later years of births averted in the early days of a program. Obviously, if we ignore these extra later births, the impact of a program is overestimated. If only those births not made up in later years are counted as averted, we underestimate the effect, since postponement of births increases the mean age of childbearing, thereby reducing the growth rate even

([3]) See Chapter 6.

if every cohort still bears the same number of children. Chandrasekaran and his colleagues therefore propose that an estimate of expected births in the year following the start of a program be derived from a reproductive model under two sets of assumptions:

i) no family planning, and
ii) family planning as introduced by the actual program.

Births averted in a given year t would be estimated by differencing the results from (i) and (ii) for that year. During the period of actual observation, the results from (ii) should be close to the observed number of births.

Although I know of no application of this technique, it is one that deserves attention in the future, always bearing in mind that projection or extrapolation techniques are accurate only insofar as they are based on accurate assumptions about future trends in reproductive behavior.

IV. DISCUSSION AND FUTURE RESEARCH

Simulation models are powerful tools in the study of demographic processes. Their utility has been demonstrated; their potential contribution is great. Just a few possible projects can be mentioned here.

In the area of estimation and measurement, further investigations of the effect of different levels of demographic determinants could enable us to group variables in two categories: those in which small variations in level or distribution have a relatively large effect on dependent variables (e.g., birth rates) and those to which the dependent variables are relatively insensitive. Precise estimates of the first category of variable will be necessary for simulation and study of a real population, whereas a rough approximation may serve for the second group. One result of this kind of study would be to guide the choice of data to be collected prior to the introduction of a family planning program—to serve as a basis for selecting a program and for later evaluation of its impact.

Another area is target setting. A population birth rate can be reduced by many different types of programs. Simulation of alternatives as to eligibility criteria, contraceptive methods to be used, acceptance patterns, switching patterns, etc. can give estimates of the numbers of acceptors in different categories and at different stages during the program required for the achievement of the desired

decline in births. Other studies could relate the number and characteristics of acceptors to costs, personnel and supplies necessary to attract women into the program, thereby offering planners both increased awareness of the kinds of programs which could be instituted, and a way of evaluating them and deciding which alternatives meet their needs.

Potter (1972) has proposed that estimates of births averted in successive generations be attempted via simulation. As discussed previously, current methods consider only the births to a cohort. However, as family planning continues, it affects population growth in a second way. Fewer women enter the childbearing ages in the next generation than would have done so in the absence of a program and they reproduce at a lower rate, thus averting further births. Simulation procedures could estimate this two-pronged effect under vigorous assumptions about the permanency of various changes in reproductive behavior.

Methodological studies also deserve concerted attention. In particular, simulation of data subject to response errors can help in detecting those indices least responsive to a variety of such errors. However, data on the nature and patterning of various types of errors would be required.

Another application, one which may become possible in the future as the ability to "fit" real data by models improves, is in assessing the impact of a program when evidence exists that changes in reproductive behavior began prior to the introduction of family planning. Two questions plague evaluators. Would the observed declines in fertility have occurred even without the program? If not, what is the magnitude of the program impact? Computer simulation, enabling investigators to carry out many calculations rapidly and relatively inexpensively, could become an important tool for attacking this problem. The population could be simulated under different sets of assumptions about the continued trend in reproductive parameters (e.g., no further change, same rate of change, increased change at various rates), and the results compared with observed values. The differences represent estimates of program effect. Although this procedure does not offer an investigator a single answer to his question, it does provide a basis for more informed judgment.

The formulation of a model, to return to some of the more general aspects of this technique, is in itself valuable. While verbal theories may express ideas well, it is frequently difficult to determine precisely their assumptions about variables and interrelationships. A simulation model entails rigorous and exact definitions of variables and statements about their form and dependencies. It also indicates data

requirements and draws attention to areas where available data are deficient.

Different types of models are appropriate to different types of problems. A biological approach to fertility is a necessity if detailed studies of the effects on fertility indices of changing contraceptive effectiveness, abortion rates, and multiple method family planning programs are to be carried out. A demographic approach may be more appropriate to investigate intergenerational effects of changes in birth rates.

A cross-sectional population model is a more convenient tool for examining period changes in reproductive parameters. A similar study with a cohort model would require several simulations, introducing the changes at different ages in each one, and then combining the output in some way.

Two-sex models are necessary for the study of couple-formation and its effect on fertility; however, this type of model has proved difficult to develop.

It seems trite to say that people vary on any dimension imaginable, and therefore a model which allows probabilities to vary among individuals is more realistic than one that does not. Two issues are discernible here. First, a heterogeneous population consists of groups (where the group size may be as small as one) of homogeneous individuals. The study of the generally more restricted homogeneous population models may lead more easily to insights and conclusions for specific values which can later be combined into conclusions about the more realistic heterogeneous groups. Also, the issue of realism enters into any discussion (Leridon and Jacquard, 1973) of models

> "in relation to many features: the structure of the model population, the rules for generating events and changing the population, the relationships assumed and, where applicable, the numerical values used. Since any model entails simplification of reality, this issue often amounts to a judgment as to which aspects of reality may be neglected or simplified without unduly affecting the conclusions. Such judgments are necessarily related to the resources and data available and to the purpose of an application. Often judgments on this score also have a flavor of personal idiosyncrasy: individual investigators vary as to which particular simplifications they find acceptable and which outrageous" (Sheps, 1971).

One object of model studies has been to treat just this problem: to examine the consequences of making a less complex assumption by comparison of results of simulations with and without this premise, and to determine the bias which results from a certain simplification (cf. Barrett, 1971; Holmberg, 1972).

The various simulation methods each offer both advantages and drawbacks. Analytical models are restricted and difficult to solve. Often results are applicable only to large populations or after a long period of constant parameter values. On the other hand, when an analytical solution is reached, it offers results for the whole range of values of the variables. Macro- and microsimulation are numerical techniques and results can be generalized only by testing a range of values, observing the consequences and reaching "empirical" conclusions.

Sheps (1971) has rightly observed that:

> "The construction of a computer model, especially microsimulation, is time-consuming and expensive. After it is constructed, it may be too easy to test different combinations of parameters on a completed model, with a rapid accumulation of masses of tables that may never be fully analyzed. Efficient use of completed models requires careful planning. A computer model sometimes tempts investigators not to think enough about their problems. Questions that can be explored much more suitably by mathematical analysis may instead be automatically submitted to the computer.
> "Computer models have valuable heuristic uses. Even when the computer simply provides numerical illustrations of results from analytic models, the implications of these results may appear clearly only from such examples. Unexpected results of a simulation experiment may bring a new insight into problems and a more profound understanding of relationships, and may stimulate analytic investigations and new approaches to empirical observations."

Computer models are also the only ones with which investigations of reactions to changes in parameters in successive time periods have been feasible.

Macrosimulation models are usually simpler to develop and yield exact expectations but are limited by computer capacity to fewer variables and restricted output.

Microsimulation models, in contrast, may be more flexible and realistic. A wide range of hypotheses about levels and interactions of variables can be studied and reasonable limits set for values of variables that are unobservable or for which data are poor. Because their output represents a sample from a population, variation in results between samples and as a function of sample size can be considered. Changes with calendar time and studies of family planning acceptance patterns require models with these provisions.

In general, for a particular question, the simplest model capable of handling it is the best choice. The exception here is that the simplest and least expensive procedure for studying almost any facet of fertility via simulation may be to employ an already existing model, one designed as a flexible tool for many research purposes. To this

end, users' manuals for at least two of the large-scale microsimulation models, POPSIM (U.S. National Center for Health Statistics, 1973) and REPSIM have been prepared and will be readily available.

The application of simulation to the determination and evaluation of family planning programs is just beginning. So long as we recognize that a model is the extrapolation of a set of assumptions about reality rather than reality itself, the investigation of the foregoing problems with existing and forthcoming models shows great promise for the future.

APPENDIX: EXISTING MODELS

ANALYTICAL MODELS

Much of the work on analytical models based on a demographic approach to fertility has been summarized by Keyfitz (1968) and Coale (1972). The biological approach has been pursued and reviewed in Henry (1972), Sheps, Menken and Radick (1969) and Sheps and Menken (1973).

Although analytical models have provided penetrating insights into problems of family planning programs, they include relatively few variables and quite restrictive assumptions and will not be discussed further here.

MICROSIMULATION MODELS WITH A DEMOGRAPHIC APPROACH TO FERTILITY

Examples of macro- and microsimulation models will be reviewed briefly. Unfortunately, descriptions of the various models are usually not detailed enough to determine precisely the extent of similarities or differences between the more complex models. Macrosimulation models which formulate fertility demographically are actually population projection models and will not be discussed here.

POPSIM, described in section III, is a flexibly designed microsimulation model. Studies, many of them unpublished, include the effects of changing marriage patterns, and of alternative family planning programs on fertility and a number of methodological investigations.

Orcutt and his colleagues (1970) are attempting to develop a model of income distribution by simulating the effect of the behavior of family members on the family's income status over time. For this purpose, they are developing a microsimulation model with demographic components. A long-term goal is to provide links between

individual probabilities and the state of the current economy, e.g., to have desired number of children dependent upon overall economic conditions. Their model is still in its early stages of development.

MACROSIMULATION MODELS WITH A BIOLOGICAL APPROACH TO FERTILITY

FERMOD, originated by Potter and his associates, was also discussed in section III. More recently, Potter (1971) has developed ACCOFERT, a model which projects forward in time a cohort of women of the same age who have just delivered and who have accepted a method of family planning. Fecundability and sterility are assumed to be functions of age, as is the probability of a fetal loss or a stillbirth. Duration of pregnancy and of the anovulatory period following an abortion are both constant. The postpartum nonsusceptible periods following a stillbirth or a live birth have a probability distribution. Contraceptive effectiveness is constant and there is a discontinuation life table. Work with this model has quantified the inadequacy of a family planning program limited to a single contraceptive method and with a significant drop-out rate. It has also been used to study births averted by various methods.

A model was developed at the University of Cairo (Tolba, 1966) to supply numerical results for a version of the Perrin-Sheps (1964) analytical model. All parameters (fecundability, probability of fetal loss, durations of pregnancy and postpartum nonsusceptibility) are constant. Extensive tables have been published studying the effects of changes in the levels of these parameters, e.g., of decreasing fecundability, increasing fetal wastage, and increasing the postpartum nonsusceptible period following a live birth.

Venkatacharya (1969, 1972) has used the simulation of the mean and variance of intervals between successive births and since the most recent birth to study the effects of specified contraceptive programs and to evaluate the adequacy of the open birth interval as a measure of recent changes in fertility.

The capacity of the computer so limits the number of variables possible in a macrosimulation model that none of those above include provisions for marriage or mortality, and little variation over time or between individuals in the population is feasible.

MICROSIMULATION MODELS WITH A BIOLOGICAL APPROACH TO FERTILITY

Applications of REPSIM, not all of which have been published so far, have included: attempts to approximate empirical data (Ridley

and Sheps, 1966; Ridley *et al.,* 1969), studies of the variance of simulated results, of the effects and interactions of the specified biological and demographic factors (Sheps *et al.,* 1967), simulation of biases that may be observed in empirical data (Sheps *et al.,* 1970) and of the effects of possible contraceptive programs.

Barrett (1971, 1972) described a model of a marriage cohort in which fecundability varies among women according to a beta distribution. Sterility is a function of age. Probabilities of fetal wastage are age-dependent. The duration of pregnancy is fixed but the postpartum nonsusceptible period has a variable duration specific for each pregnancy outcome. Barrett has studied the types of data that would be needed to produce results comparable to empirical data, the timing of averted births and the pattern of approach to a new birth rate level following a change in reproductive behavior.

Jacquard (1967) implemented a model simulating a cohort of women from age 15 to 45. Probabilities of death, marriage, and following marriage, of widowhood and divorce are age-dependent. Fecundability increases from age 15 to some age a_1, plateaus for a period and then declines with age. The probability of a live birth is constant. Postpartum nonsusceptibility is a random variable depending upon fetal loss. Provisions for contraceptive effectiveness and use for spacing or limiting pregnancies are included. Length of use depends upon parity. Applications have included the study of levels of contraceptive effectiveness and an attempt (Bodmer and Jacquard, 1968), to "fit" data obtained from a group of highly fertile women.

Holmberg (1970, 1972) has reported a model in which a birth cohort of women is followed from age 15 to menopause. Each woman is exposed to age-specific probabilities of death, sterility and marriage until marriage occurs. Fecundability varies with age and its peak varies between women according to a beta distribution. Contraception of fixed effectiveness may be accepted either after the desired number of children are born or for spacing purposes. Available literature on estimates of the distributions of components of the reproductive process has been reviewed to find realistic input data. The model has been used in attempts to reproduce empirical data, and in studies both of the importance of changes in levels of different variables and of the effects of introducing additional complexity into a model.

POPREP (1973) is a modification of POPSIM that incorporates many of the features of REPSIM, so that biological parameters are included. POPREP has been applied in the study of bias in suggested fertility measures, in the testing of new measurement proce-

dures (Poole, 1973) and in experiments simulating the effect of recall error on birth intervals.

Venkatacharya (1970) examined the effects of the proportion of susceptible women on indices used in evaluating population control programs. His model follows a cohort of women, starting from a given age between 15 and 34 until they reach age 45. Fecundability varies with age. Probabilities of fetal wastage are constant. Postpartum nonsusceptibility varies according to a probability distribution.

The models considered here vary in the required form of input data. For example, Jacquard (1967) and Barrett (1971) both allow the postpartum nonsusceptible period to follow a random distribution. In the former case, the input data must specify the probability that the anovulatory period will last one month, two months, etc. Barrett, in contrast, assumes that this period is equal to one month plus the sum of two random variables, each with a geometric distribution with parameter r. Therefore, the probability of a three-month nonsusceptible period is r^2; of 4 months, $2r^2(1 - r)$, etc. The only input, then, is the value of the parameter r. Assuming that the distribution of a variable has some functional form, as in Barrett's case, usually saves both computer storage space and calculation time. The user of such a model must weigh whether or not the particular form specified is consistent with his own conclusions about the shape of the distribution.

Unfortunately, limitations of space preclude a comparison of models according to the precise assumptions about the variables included. The interested reader should see the references for the particular models.

REFERENCES

Barrett, J. C.

1971 "Use of a fertility simulation model to refine measurement technique." Demography, Vol. 8 (Nov.): 481-490.

1972 "The time response in averted births." Population Studies, Vol. XXVI (Nov.): 507-514.

Bodmer, W. F. and A. Jacquard

1968 "La variance de la dimension des familles selon divers facteurs de la fécondité." Population, Vol. 23 (Sept.-Oct.): 869-878.

Chandrasekaran, C., D. V. R. Murty and K. Srinivasan

1971 "Some problems in determining the number of acceptors needed in a family planning programme to achieve a specified reduction in the birth rate." Population Studies, Vol. 25 (July): 303-308.

Clague, Alice and Jeanne Clare Ridley

1973 "The assessment of three methods of estimating births averted." *In* Bennett Dyke and Jeanne Macluer (eds.), Computer Simulation in Human Population Studies. New York: Seminar Press.

Coale, A. J.

1972 The Growth and Structure of Human Populations. Princeton: Princeton University Press.

Davis, Kingsley and Judith Blake

1956 "Social structure and fertility: an analytic framework." Economic Development and Cultural Change, Vol. 4 (No. 3): 211-238.

Henry, Louis

1972 On the Measurement of Human Fertility: Selected Writings. Translated by Mindel C. Sheps and Evelyne Lapierre-Adamcyk. New York: Elsevier Publishing Company.

Holmberg, Ingvar

1970 Fecundity, Fertility and Family Planning I, Demographic Institute Reports 10. Gothenburg: University of Gothenburg.

1972 Fecundity, Fertility, and Family Planning II, Demographic Institute Reports 11. Gothenburg: University of Gothenburg.

Horvitz, D. G., F. G. Giesbrecht, B. V. Shah and P. A. Lachenbruch

1971 "POPSIM, A demographic microsimulation model." *In* Carolina Population Center Monograph 12. Chapel Hill: University of North Carolina, 45-63.

Jacquard, Albert

1967 "La reproduction humaine en régime malthusien." Population, Vol. 22 (Sept.-Oct.): 897-920.

Keyfitz, Nathan

1968 Introduction to the Mathematics of Population. Reading: Addison-Wesley.

Leridon, H. and A. Jacquard

1973 "How complicated should a model be?" *In* Bennett Dyke and Jeanne Macluer (eds.), Computer Simulation in Human Population Studies, New York: Seminar Press.

Orcutt, Guy H., Gerald E. Peabody and Steven B. Caldwell

1970 "Microanalytic simulation: A demographic model," 1970 Social Statistics Proceedings of the American Statistical Association: 122-128.

Perrin, Edward B. and Mindel C. Sheps

1964 "Human reproduction: A stochastic process." Biometrics, Vol. 20 (Mar.): 28-45.

Poole, Kenneth

1973 "Fertility measures based on birth interval data." Theoretical Population Biology.

Potter, Robert G.

1971 "Inadequacy of a one-method family-planning program." Social Biology, Vol. 18 (March): 1-9.

1972 "Contraceptive impact over several generations." *In* T. N. E. Greville (ed.), Population Dynamics. New York: Academic Press, Inc.; 137-166.

Potter, R. G. and G. Stephan Masnick

1971 "The contraceptive potential of early versus delayed insertion of the intrauterine device." Demography, Vol. 8: 507-517.

Potter, R. G., G. Stephan Masnick and Murray Gendell

1973 "Postamenorrheic versus post-partum strategies of contraception." Demography, Vol. 10: 99-112.

Potter, R. G. and J. M. Sakoda

1966 "A computer model of family building based on expected values," Demography, Vol. 3 (No. 2): 450-61.

Potter, R. G., J. M. Sakoda, W. E. Feinberg

1968 "Variable fecundability and the timing of births." Eugenics Quarterly, Vol. 15 (Sept.): 155-163.

Ridley, Jeanne Clare and Mindel C. Sheps

1966 "An analytic simulation model for human reproduction with demographic and biological components." Population Studies, Vol. XIX (Mar.):297-310.

Ridley, J. C., M. C. Sheps, J. W. Lingner and J. A. Menken

1969 "On the apparent subfecundity of non-family planners." Social Biology, Vol. 16: 24-28.

Sarma, R. S. S.

1970 Alternative Family Planning Strategies for India: A Simulation Experiment. Institute of Statistics Mimeo Series No. 720. Chapel Hill: University of North Carolina.

Sheps, Mindel C.

1971 "A review of models for population change." Review of the International Statistical Institute, Vol. 39 (No. 2): 186-196.

Sheps, Mindel C. and Jane Menken

1973 Mathematical Models of Conception and Birth. Chicago: University of Chicago Press.

Sheps, Mindel C. and Edward B. Perrin

1963 "Changes in birth rates as a function of contraceptive effectiveness: Some applications of a stochastic model." American Journal of Public Health, Vol. 53 (July): 1031-1046.

Sheps, Mindel C., Jane Menken and Annette Radick

1969 "Probability models for family building: An analytic review." Demography, Vol. 6 (May): 161-183.

Sheps, M. C., J. A. Menken, J. C. Ridley and J. W. Lingner

1970 "The truncation effect in closed and open birth interval data." Journal of the American Statistical Association, Vol. 65 (June): 678-93.

Sheps, Mindel C., Jeanne Clare Ridley, and Joan W. Lingner

1966 "Effects of selected factors on natality: Quantitative estimation through simulation," Proceedings on Simulation in Business and Public Health. New York: American Statistical Association, and Public Health Association of New York City, 176-188.

Tolba, Salah El-din

1966 "On estimating the effects of changing the values of some factors of fertility," Proceedings of the Institute of Statistical Studies and Research, Vol. 1. Cairo: University of Cairo Press: 81-178.

U.S. National Center for Health Statistics

1973 Users' Manual for POPSIM, DHEW Publication No. (HSM 73-1216). Rockville, Md. (June).

Venkatacharya, K.

1969 "Certain implications of short marital durations in the analysis of live birth intervals." Sankhya, Vol. 31B (June): 53-68.

1970 "Some implications of susceptibility and its application in fertility evaluation models." Sankhya, Vol. 32B (June): 41-54.

1972 "Some problems in the use of open birth intervals as indicators of fertility change." Population Studies, Vol. 26 (Nov.): 495-506.

Chapter 11

THE IMPACT ON FERTILITY OF SINGAPORE'S FAMILY PLANNING PROGRAM

by
C. T. Chang and P. M. Hauser

I. INTRODUCTION

Singapore, lying 77 miles north of the equator and at the southern tip of the Malay Peninsula, is one of the smallest nations on earth. It has an area of 225 square miles, comprising the main island of Singapore and numerous islets. With a population of about 2.1 million (2,074,507 according to the 1970 census), the density exceeds 9,000 persons per square mile (1).

The history of modern Singapore dates from 1819 when Sir Stamford Raffles founded the settlement. From that time onward Singapore remained under British rule until the Japanese invaded the island during the Second World War. It was returned to the British after the war, but became a self-governing state in 1959. In 1963 Singapore joined the Federation of Malaysia but the merger proved politically untenable and, as a result, Singapore seceded from the Federation in 1965 and became an independent state.

Although the recent demographic history of Singapore has been covered in detail elsewhere (Saw Swee-Hock, 1970a and You Poh Seng *et al.*, 1971), it is worth outlining a few characteristics here. The present population of 2.1 million is the result of a more than doubling of the population since 1947 (938,144) and represents an increase of more than 40 per cent since 1957 when the population

(1) Three censuses have been taken in Singapore since the end of the Second World War, one in 1947, one in 1957 and the most recent in 1970. In addition, a national sample household survey was conducted in 1966. The census statistics may be regarded as reasonably accurate and complete, although some specific items, e.g. the data on fertility collected in the 1947 census, have been found to be unreliable. Statistics of vital events, that is of births, deaths and marriages, are compiled monthly from registration records. Birth and death statistics are of high quality; marriage statistics, however, are less complete because marriages by religious and customary rites still exist and may not always be registered.

numbered 1,445,929. Unlike the pre-Second World War period when the major source of population increase was migration, mainly from China and the Indian sub-continent, the growth of population after the war is attributable predominantly to natural increase. The annual rate of natural increase in the postwar period averaged about 35 per thousand until the early sixties, at which time it dropped to below 30.

The main ethnic groups in Singapore are the Chinese (76.2 per cent of the total population in 1970), Malays (15.0 per cent) and Indians and Pakistanis (7.0 per cent). The remaining 1.8 per cent comprises Eurasians, Europeans and others.

Since immigration was until recently the major source of population increase, the population had been characterized by an exceedingly high sex ratio. In 1860, when the sex ratio reached its peak, there were on the average six males for each female in the population. Since then, the sex ratio has continuously decreased but it was not until after the war that the two sexes became approximately equal in numbers. The sex ratio was 1.22 in 1947, 1.12 in 1957, and 1.05 in 1970. This more normal sex ratio was, however, an average over considerable ethnic variation and although in 1970 the sex ratio was 1.02 for the Chinese and 1.04 for the Malays, it was substantially higher at 1.52 for Indians and Pakistanis.

Owing to the high rates of natural increase after the Second World War, the population has been relatively young in recent years. In 1957, 41.9 per cent of the male and 43.8 per cent of the female population were under age 15. These proportions had decreased slightly to 38.9 per cent for males and 38.7 per cent for females by 1970. The Malays had a younger age-structure than the Chinese and although the Indian and Pakistani males were on average older than the Chinese, the proportion of their females under 15 was as large as among the Malays. This sex difference in age composition can be traced to past immigration which was heavily male.

Singapore has experienced relatively rapid social and economic development since the end of the Second World War. By 1970, only 11.5 per cent of the men and 40.1 per cent of the women were without formal education. The educational level of women was rapidly improving as indicated by the fact that 58.7 per cent of females aged 10 to 19 had completed primary school and 33.3 per cent secondary school, compared with 38.2 per cent and 20.7 per cent respectively for women aged 20-34.

The higher educational levels of the younger women are accompanied by higher rates of economic activity. Between 1957 and 1966, the labor force participation rate increased from 22.9 to 40.9 per cent for women aged 20-24 and from 16.8 to 23.6 per cent for women

aged 25 to 34 years old. Both the higher educational level of women and their increasing economic activity rates have undoubtedly contributed to the rising age at marriage described below.

Important changes have occurred in the economic structure in recent years. The most prominent of these is the emergence of the manufacturing sector as the new base for the economy. Whereas manufacturing and quarrying accounted for only 9.1 per cent of Gross Domestic Product (GDP) in 1960, it accounted for 20.7 per cent in 1970. In line with the industrialization of the economy, the proportion of the labor force employed in manufacturing increased from 15.6 to 19.3 per cent between 1957 and 1966. On the other hand, the proportion engaged in agriculture and fishing dropped from 7.0 to 3.5 per cent. As a result of the rapid growth of the economy, per capita GDP more than doubled in a decade from Singapore $1,246 in 1959 to Singapore $2,716 in 1970 (Lim Chong Yah and Ow Chwee Huay, 1971).

At the same time, great improvements have been made in the housing conditions of the population. Two five-year public housing programs were developed in the sixties in an effort to tackle the problem of overcrowded and deteriorating housing and a third one is now being implemented, covering the period from 1970 to 1975. These massive programs have succeeded in housing two-fifths of the population in high-rise public housing apartments.

As has been noted by You Poh Seng (1967) and others (You Poh Seng *et al.*, 1971), these social and economic changes, together with improvements in health standards and changes in the age composition, have led to a rapid decline in mortality. The crude death rate dropped from 13.3 per thousand in 1947 to 7.4 per thousand in 1957 and has remained between 5.0 and 5.5 per thousand since 1965. The reduction in infant mortality has been even more remarkable; from a level of 87.3 per thousand in 1947, the infant mortality rate per thousand declined to 41.4 in 1957 and 20.5 in 1970.

The declining death rate after the Second World War, coupled with a birth rate which remained relatively high until the mid-fifties, generated, of course, a great rise in the rate of natural increase—a pattern consistent with the classical description of the demographic transition. Since 1957 the birth rate has declined from a level of 42.7 per thousand to 22.3 per thousand in 1971. The following analysis explores the factors accounting for the fertility decline in Singapore against the background of continued rapid modernization.

FERTILITY CHANGES IN THE POSTWAR PERIOD

1. *BACKGROUND.* The fertility of the population had been rising prior to the outbreak of the Second World War. From a level of 21 per thousand at the turn of the century, the crude birth rate had increased to about 30 per thousand in the early twenties and had reached 46 per thousand on the eve of the war. Part of this upward trend was attributable to the increasing balance between the sexes but part was a reflection of a genuine rise of fertility in the period. The general fertility rate and other fertility measures had all increased continuously before the war, independently of the changing sex ratio of the population.

The upward trend in fertility had characterized all three major ethnic groups, but especially the Chinese and the Indians and

Table 11.1. *Crude birth rate by ethnic group, 1947-1971* (a)

Year	Total population	Chinese		Malays		Indians and Pakistanis	
		(1)	(2)	(1)	(2)	(1)	(2)
1947	45.9	46.1	(b)	48.1	(b)	44.8	(b)
1952	45.4	45.3	(b)	47.1	(b)	41.5	(b)
1957	42.7	42.4	42.6	47.3	49.8	40.5	35.9
1958	41.1	40.2	40.3	47.9	49.8	40.7	36.8
1959	39.4	38.0	38.2	47.4	49.2	40.1	36.3
1960	37.5	35.9	36.1	45.6	47.4	39.8	35.8
1961	35.2	33.3	33.4	45.3	46.9	36.5	33.0
1962	33.7	31.5	31.7	45.0	46.5	35.3	31.8
1963	33.2	30.8	31.0	45.0	46.2	35.0	31.9
1964	31.6	29.1	29.3	43.9	45.0	34.4	31.7
1965	29.5	27.3	27.6	39.7	40.7	32.6	29.8
1966	28.3	26.0	26.2	38.2	38.9	32.0	29.9
1967	25.6	23.8	24.1	33.3	33.7	28.1	26.4
1968	23.5	22.4	22.5	28.8	29.2	24.0	23.0
1969	21.8	21.1	21.3	24.3	24.5	23.4	22.5
1970	22.1	21.8	22.0	23.5	24.0	21.9	19.8
1971	22.3	22.2	22.4	22.8	23.2	20.9	19.9

(a) From 1953 the figures for live births are by year of occurrence; for earlier years they are by year of registration. Beginning in 1957, the figures for live births exclude births to wives of non-locally domiciled services personnel.

(b) Not available.

(1) Based on statistics of live births classified by ethnic group of *father.*

(2) Based on statistics of live births classified by ethnic group of *mother.*

Source: Singapore, Report on the Registration of Births, Deaths and Marriages, 1971.

Pakistanis. The general fertility rate of the Chinese rose from 139 to 202 per thousand in the two decades between 1911 and 1931, and that of the Indians and Pakistanis from 164 to 292 during the same period. The fertility rate of the Malays increased from 149 to 165 per thousand.

The war years witnessed a considerable dip in the upward trend of the crude birth rate for both the Chinese and the Malays and consequently for the total population. The crude birth rate of the Indians and Pakistanis had been only moderately high before the war, by reason of their high sex ratio, and its drop during the war was relatively small.

2. *POSTWAR FERTILITY CHANGES.* As in many other countries, Singapore experienced a postwar "baby boom". When the first postwar census was taken in 1947, the crude birth rate of the population was found to have jumped from its wartime level of 36.5 per thousand to 45.9 per thousand. Throughout the following decade, the crude birth rate remained high, fluctuating above the level of 45 per thousand until the mid-fifties, and it is only since 1957 that it has persistently declined. Once the downward trend set in, however, the pace of decline was great. Whereas in 1956 the crude birth rate was still high at 44.4, by 1965 it had dropped to below 30. The birth rate decreased further to 21.8 in 1969, and then rose slightly to 22.1 in 1970 and 22.3 in 1971. Thus, in about a decade and a half, the crude birth rate of the population had been reduced by about 50 per cent (Table 11.1).

Although the crude birth rate had remained largely constant before the mid-fifties, the fertility of women aged under 20 did drop by about one-fourth between 1947 and 1957. As will be observed presently, this was attributable to increased age at marriage, a postwar development reversing the prewar trend to marry younger. The total fertility rate changed little, being 6,465 per thousand in 1957 compared with 6,566 per thousand in 1947. It was only after 1957 that age-specific fertility began to decline significantly and at an accelerating rate. The fertility decline in the earlier periods was greater among younger women and in the later periods among older women (see Table 11.2). In the three-year period 1957 to 1960 the total fertility rate decreased by 9.2 per cent; between 1960 and 1965, by 18.5 per cent and between 1965 and 1970, by 35.1 per cent. By 1970 the total fertility rate had dropped to 3,104, a level less than half the 1957 figure.

3. *ETHNIC DIFFERENCES IN FERTILITY DECLINE.* The pattern of fertility decline has varied among the three ethnic groups, but

Table 11.2. *The total population: Age-specific fertility rates, 1947-1970*

Age group	Age-specific fertility rates					Percentage of change			
	1947	1957	1960	1965	1970	1947-1957	1957-1960	1960-1965	1965-1970
15-19	102.0	77.6	62.6	36.5	26.2	−24	−19	−42	−28
20-24	315.0	297.5	268.6	220.7	139.5	−6	−10	−18	−37
25-29	334.8	348.6	320.3	266.5	209.6	4	−8	−17	−21
30-34	270.5	284.6	264.1	213.7	138.5	5	−7	−19	−35
35-39	196.8	192.1	174.5	156.4	74.8	−2	−9	−10	−52
40-44	83.5	80.8	71.7	55.3	27.4	−3	−11	−23	−51
45-49	10.6	11.8	11.8	8.1	4.7	11	0	−31	−42
TFR (a)	6,566	6,465	5,868	4,786	3,104	−1.5	−9.2	−18.5	−35.1

(a) Total fertility rate.

Note: The population bases of rates for 1960 and 1965 are derived by interpolation from population estimates by Saw Swee-Hock in "Estimation of Singapore's population by quinary age group for 1962 and 1967," *The Malayan Economic Review*, XV: 1 April 1970, pp. 101-105. However, the estimates have been first revised to tally with new intercensal estimates of the total population and ethnic subpopulations made by the government Department of Statistics after the 1970 census.

Source: C. T. Chang, Fertility Transition in Singapore, Singapore University Press, 1974.

Table 11.3. *Chinese population: Age-specific fertility rates, 1957-70*

Age group	Age-specific fertility rates				Percentage of change		
	1957	1960	1965	1970	1957-1960	1960-1965	1965-1970
15-19	46.8	38.2	22.9	21.0	−18	−40	−8
20-24	282.1	232.0	184.0	128.3	−18	−21	−30
25-29	361.0	324.2	250.9	216.1	−10	−23	−14
30-34	308.4	272.1	209.1	141.4	−12	−23	−32
35-39	210.0	184.0	134.6	70.5	−12	−27	−48
40-44	88.5	78.7	58.1	25.3	−11	−26	−56
45-49	12.6	12.1	7.9	4.4	−4	−35	−44
TFR (a)	6,547	5,707	4,338	3,035	−12.8	−24.0	−30.0

(a) Total fertility rate.

Note: The population bases of rates for 1960 and 1965 are derived by interpolation from population estimates by Saw Swee-Hock in "Estimation of Singapore's population by quinary age group for 1962 and 1967," *The Malayan Economic Review*, XV: 1 April 1970, pp. 101-105. However, the estimates have been first revised to tally with new intercensal estimates of the total population and ethnic subpopulations made by the government Department of Statistics after the 1970 census.

Table 11.4. *Malay population: Age-specific fertility rates, 1957-70*

Age group	Age-specific fertility rates				Percentage of change		
	1957	1960	1965	1970	1957-1960	1960-1965	1965-1970
15-19	194.0	166.1	91.3	46.8	−14	−45	−49
20-24	337.8	384.0	378.9	180.0	14	−1	−52
25-29	317.4	309.4	308.3	186.1	−3	−0.4	−40
30-34	221.4	246.3	242.4	134.4	11	−2	−45
35-39	132.6	155.9	166.1	104.4	18	7	−37
40-44	41.6	47.1	55.1	40.5	13	17	−26
45-49	10.4	10.3	10.5	7.9	−1	2	−25
TFR (a)	6,276	6,596	6,263	3,501	5.1	−5.0	−44.0

(a) Total fertility rate.

Note: The population bases of the rates for 1960 and 1965 are derived by interpolation from population estimates by Saw Swee-Hock in "Estimation of Singapore's population by quinary age group for 1962 and 1967," *The Malayan Economic Review*, XV: 1 April 1970, pp. 101-105. However, the estimates have been first revised to tally with new intercensal estimates of the total population and ethnic subpopulations made by the government Department of Statistics after the 1970 census.

Table 11.5. *Indian and Pakistani population: Age-specific fertility rates, 1957-70*

Age group	Age-specific fertility rates				Percentage of change		
	1957	1960	1965	1970	1957-1960	1960-1965	1965-1970
15-19	243.6	193.3	99.7	44.0	−21	−48	−56
20-24	392.6	495.8	462.3	204.3	26	−7	−56
25-29	340.5	336.4	362.7	190.5	−1	8	−47
30-34	255.7	263.3	221.2	120.0	3	16	−46
35-39	154.7	149.3	111.1	59.2	−3	−26	−47
40-44	55.5	34.8	34.1	18.8	−37	−2	−45
45-49	4.8	14.6	7.9	1.6	204	−46	−80
TFR (a)	7,237	7,438	6,495	3,192	2.8	−12.7	−50.9

(a) Total fertility rate.

Note: The population bases of the rates for 1960 and 1965 are derived by interpolation from population estimates by Saw Swee-Hock in "Estimation of Singapore's population by quinary age group for 1962 and 1967," *The Malayan Economic Review*, XV: 1 April 1970, pp. 101-105. However, the estimates have been first revised to tally with new intercensal estimates of the total population and ethnic subpopulations made by the government Department of Statistics after the 1970 census.

the difference has been especially great between the Chinese and the other two groups. As the predominant ethnic group, constituting three-fourths of the total population, the Chinese pattern of fertility changes has essentially been that of the total population. Although the crude birth rates of the other two groups also started to decline at about the same time in the mid-fifties, their total fertility rates increased between 1957 and 1960 (2) and it was not until the period 1960-65 that the total fertility rates of these smaller ethnic groups also began to decrease. Their rate of decrease accelerated so that although in 1965 their total fertility rates were substantially higher than that of the Chinese, by 1970 the ethnic differences had been much reduced (Tables 11.3-11.5).

4. *AGE-PATTERN OF FERTILITY.* As a result of the differential changes in age-specific fertility in Singapore, its pattern has greatly changed (3). The relative contribution of women under 25 to the total fertility of the population has persistently decreased since 1947, and so has fertility at ages above 40. The relative contribution of women aged 35-39 has also shown a considerable drop in the period 1965-70. On the other hand, the contribution to total fertility of women at ages 25-29 has increased continuously from 25.5 per cent in 1947 to 33.8 per cent in 1970. Consequently, the age pattern of fertility has shifted from a broad-peak type in which fertility at ages 20-24 and 25-29, while higher than at other ages, differs little from each other, to a late-peak type in which fertility in the late twenties considerably exceeds fertility in other age groups. Thus, in this process of transition the total fertility of the population has become more and more narrowly concentrated in the age range 25-34. Whereas in 1947, 46.1 per cent of total fertility resulted from child-bearing to women in this age interval, the proportion had increased to 56.1 per cent by 1970 (Table 11.6).

Among the Chinese the relative contribution to total fertility of women aged under 25 changed little between 1957 and 1970. As was the case for the total population, however, the fertility contribution of Chinese women aged 25-29 increased substantially from 27.6 to 35.6 per cent. On the other hand, fertility at ages over 35 became

(2) Because of inter-ethnic marriages the number of live births classified by ethnicity of *father* differs from the number classified by *mother.* Since only statistics on births by ethnicity of mother are available for computing age-specific rates, crude birth rates by ethnicity of mother should be used if changing birth rates are to be related to changes in age-specific fertility rates and hence to total fertility rates.

(3) The different age-patterns of fertility have been the subject of a study by the United Nations (1965).

Table 11.6. *The total population: Relative contribution of each age group to the total fertility rate, 1947-70 (in percentages)*

Age group	*1947*	*1957*	*1965*	*1970*
15-19	7.8	6.0	3.8	4.2
20-24	24.0	23.0	23.1	22.5
25-29	25.5	27.0	27.8	33.8
30-34	20.6	22.0	22.3	22.3
35-39	15.0	14.9	16.3	12.0
40-44	6.4	6.2	5.8	4.4
45-49	0.8	0.9	0.9	0.8
Total	100.1 (a)	100.0	100.0	100.0

(a) Due to rounding.
Source: Based on Table 11.2.

Table 11.7. *Relative contribution of each age group to the total fertility rate, by ethnic group, 1957-70 (in percentages)*

Age group	*Chinese*			*Malays*			*Indians and Pakistanis*		
	1957	*1965*	*1970*	*1957*	*1965*	*1970*	*1957*	*1965*	*1970*
15-19	3.6	2.6	3.5	15.5	7.3	6.7	16.8	7.7	6.9
20-24	21.5	21.2	21.1	26.9	30.2	25.7	27.1	35.6	32.0
25-29	27.6	28.9	35.6	25.3	24.6	26.6	23.5	27.9	29.8
30-34	23.6	24.1	23.3	17.6	19.4	19.2	17.7	17.0	18.8
35-39	16.0	15.5	11.6	10.6	13.3	14.9	10.7	8.6	9.3
40-44	6.8	6.7	4.2	3.3	4.4	5.8	3.8	2.6	2.9
45-49	1.0	0.9	0.7	0.8	0.8	1.1	0.3	0.6	0.3
Total	100.1 (a)	99.9 (a)	100.0	100.0	100.0	100.0	99.9 (a)	100.0	100.0

(a) Due to rounding.
Source: Based on Tables 11.3-11.5.

much less important, declining from 23.8 per cent of the total fertility in 1957 to 16.5 per cent in 1970 (Table 11.7).

For the Malays a striking feature of the changing age pattern in fertility concerns women at ages above 35, whose contribution to total fertility increased from 14.7 per cent in 1957 to 21.8 per cent in 1970. On the other hand, childbearing at ages under 20 decreased its contribution to the total fertility of the group, declining from 15.5 to 6.7 per cent.

As among the Chinese, the fertility of the Indians and Pakistanis at ages 25-29 showed a continuous increase in relative importance over the years, but unlike the Chinese and like the Malays, their fertility at ages under 20 became less important in 1970 than in 1957. The relative contribution to total fertility of Indian and Pakistani women aged over 35 changed much less than that of either Chinese or Malay women.

5. *DIFFERENTIAL FERTILITY.* As fertility declined in Singapore ethnic fertility differentials changed significantly (Tables 11.3-11.5). In 1957, at the onset of the transition, Malays had the lowest total fertility rate and Indians and Pakistanis the highest. With the earliest fertility reduction occurring among the Chinese while fertility was rising among both the other ethnic groups, Chinese fertility, from an intermediate level in 1957, had become the lowest by 1960 and has remained so ever since.

In 1960, Indians and Pakistanis were the most fertile of the three ethnic groups. Since then, however, their fertility has declined faster than that of the Malays so that by 1970 Malay fertility was higher than that of the Indians and Pakistanis. Thus, between 1957 and 1970, Malay fertility shifted from the lowest to the highest level, while the Chinese and the Indians and Pakistanis retained their positions relative to one another.

Other fertility differentials observable in Singapore are in keeping with those in economically advanced areas.* Fertility is higher in rural than in urban areas, it is inversely related to education and to income, and varies as expected by occupation. Further details of differential fertility have been described elsewhere (Chang, 1974).

CHANGES IN NUPTIALITY

Evidence suggests that there was for a time before World War II a trend for women to marry younger. This trend to marry earlier was reversed, however, after the end of the war. Whereas in 1947 women remaining single formed 72 per cent of all women at ages 15-19

and 16.5 per cent at ages 20-24, by 1957 the proportions had increased to 80.0 per cent and 33.0 per cent respectively. The trend continued after 1957, and in 1970 more than 95 per cent of women aged 15-19 were still single. This reversal of trend occurred among all three ethnic groups, but with some time lag among the Malays. Thus, before the onset of fertility decline, age at marriage of the population was already rising, a factor potentially conducive to fertility decline.

On the other hand, at the same time as the age at marriage was rising, the risk of women becoming widowed was rapidly declining due to the decreased mortality (Table 11.9). Whereas in 1947 the proportions of widowed to married women were 18.6 per cent for women aged 35-39, 30.8 per cent at ages 40-44, and 51.8 per cent at ages 45-49, in 1966 these proportions had declined to only 4.5, 11.5, and 21.1 per cent respectively. As with the case for rising age at marriage, this continuous reduction in widowhood occurred among all three ethnic groups. Such reductions meant, of course, that a larger proportion of women were remaining married through their

Table 11.8. *Percentage of women remaining single by age, ethnic group and year*

Ethnic group and year	Age group						
	15-19	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*	*45-49*
All races							
1947	71.6	16.5	8.9	7.0	5.0	6.8	7.2
1957	80.0	33.0	9.2	4.5	4.0	4.9	5.7
1966	94.3	55.8	19.9	7.2	3.8	2.9	3.0
1970 (a)	95.2	64.6	22.6	9.6	5.1	3.3	3.1
Chinese							
1947	80.2	18.2	9.7	7.6	5.2	7.3	7.6
1957	87.4	39.6	10.5	4.7	4.3	5.2	6.2
1966	96.7	63.2	23.5	8.2	4.1	3.0	2.9
Malays							
1947	32.7	6.3	2.9	2.3	2.2	2.0	2.1
1957	46.6	9.7	3.3	1.7	1.1	1.4	0.9
1966	82.9	31.3	4.0	1.4	0.8	0.0	1.5
Indians and Pakistanis							
1947	35.7	7.6	3.0	1.5	2.0	1.4	2.4
1957	52.0	13.9	4.1	1.9	1.2	1.0	0.8
1966	84.3	29.9	5.4	6.1	1.0	2.7	1.5

(a) At the time of writing, no detailed figures for the ethnic groups are available.

Table 11.9. *Per cent widowed of ever-married women by ethnic group, age and year*

Ethnic group and year	Age group						
	15-19	25-29	30-34	35-39	40-44	45-49	20-24
All races							
1947	1.8	2.5	4.4	8.2	15.6	23.4	34.0
1957	0.5	0.8	1.7	3.5	7.7	15.4	26.6
1966	0.0	0.4	1.4	1.7	4.2	10.1	17.2
Chinese							
1947	1.2	2.4	4.5	8.4	16.6	23.7	34.8
1957	0.3	0.7	1.7	3.6	7.8	15.5	26.8
1966	0.9	0.3	1.1	1.7	3.9	10.4	16.3
Malays							
1947	2.9	3.0	4.2	8.1	11.2	23.2	30.5
1957	0.7	1.2	2.3	3.4	7.2	16.2	27.3
1966	0.0	1.0	3.1	2.1	6.2	11.5	21.7
Indians and Pakistanis							
1947	1.7	2.1	3.1	5.0	13.1	25.6	35.0
1957	0.6	0.1	1.4	2.4	7.8	15.3	28.3
1966	0.0	1.1	0.7	2.9	4.9	4.2	22.8

Source: M. V. Del Tufo, Malaya: A Report on the 1947 Census of Population, 1949.
S. C. Chua, State of Singapore: Report on the Census of Population, 1957, 1964.
You Poh Seng *et al.*, Singapore Sample Household Survey, 1966, Report No. 7, 1967.

reproductive period. For obvious reasons, this trend must have, to some extent, offset the effects of rising age at marriage on the fertility level of the population.

There have also been other nuptiality changes bearing upon the fertility of the population. These include the increasing proportions of women ever married in the later reproductive ages, perhaps reflecting the prewar trend of women marrying younger, and a slight increase in marriages ending in divorce in recent years. These changes were minor, however, and their effects on fertility have therefore been insignificant.

Although the reduction in widowhood has been quite substantial, it has affected mainly the marital patterns of women in their later reproductive ages. Its effect on fertility has therefore been relatively small compared with the fertility impact of marriage postponement. The effect of nuptiality changes on Singapore fertility will be considered in relation to other factors below.

II. THE FAMILY PLANNING PROGRAM

THE FAMILY PLANNING ASSOCIATION

The organized family planning program has had a relatively long history in Singapore. Soon after the end of the Second World War, as a consequence of social dislocation and food shortage, the Government was confronted with the problem of feeding thousands of undernourished children. The difficulties encountered in providing free meals to the children created awareness of the fact that part of the problem lay in the large numbers of children. It was concluded that to deal with the problem more effectively, parents must be assisted in planning their family size according to their means. Hence, the idea of making family planning advice available to the public was conceived. The idea was implemented by a small group of social workers and physicians who formed the Singapore Family Planning Association (SFPA) in July 1949.

At the time of its foundation, the Association was the first of its kind to appear in any country in Asia. As originally conceived, the Association's primary objective was to improve the broadly-defined welfare of families and the health of mothers and their children on an individual basis. It was soon realized, however, that the family planning program was closely tied up with the concern over the problem of rapid population growth in Singapore; but despite this, the original concern with family welfare was never abandoned, and several welfare programs having nothing to do with contraception and birth control were maintained over the years.

Although originally formed as a voluntary organization, the Association from the very beginning received an annual grant from the government; and a government policy in support of family planning activities was announced in 1959. The activities of the Association expanded quickly, and during its first year, it had 11 family planning clinics in operation, in both urban and rural areas. There were 25 clinics by 1956; and since then 25 to 30 clinics have been maintained.

From a level of about 600 acceptors in its first year of operation, 9,845 women attended the Association clinics for family planning services during 1965, its last year. In the nearly two decades of its history, the Association reached and provided family planning services to more than 88,000 women.

In the experience of the Association it was the Chinese who were the first group to respond actively to the family planning campaign, followed shortly by the Indians and Pakistanis and only later by the Malays, a pattern undoubtedly reflected in the pattern of ethnic

fertility decline. Over the years not only were larger numbers of women reached, but also the age of acceptors decreased and their number of children at time of acceptance became smaller. Moreover, although at the beginning many women came to the clinics without their husbands' knowledge, over time clinic attendance became a decision reached by both spouses. Further details about the contribution of the Association in paving the way for the government program of family planning are available elsewhere (Chang, 1974).

THE NATIONAL FAMILY PLANNING PROGRAM, 1966-70

With limited resources in the face of a rapidly expanding demand for services, the Association began to feel as early as in 1957 that the time had come for the government to take over the family planning program. The government, however, was not prepared to assume this responsibility until 1965 when, after repeated requests by the Association, it set up a review committee to consider the matter. The committee recommended that the government assume full responsibility for all family planning activities in the Republic, a recommendation that was accepted. *A White Paper on Family Planning* was published in September 1965, announcing a five-year national family planning program for the period from 1966 to 1970. The Singapore Family Planning and Population Board (SFPPB) was inaugurated shortly afterwards and was charged with the responsibility for implementing the program.

The purpose of the program was to reach 180,000, an estimated 60 per cent of all married women aged 15-44, within the five-year period. The target of 180,000 women was arrived at on the basis of a projected female population of 450,000 in the age group 15-44 by 1970, of whom 67 per cent were estimated to be married. On the further assumption that 40 per cent of the women would be either currently practicing family planning or pregnant, lactating, or sterile, and hence not currently eligible, this would leave a residual of 180,000 eligible women ([4]). It was estimated that in recruiting this number of women into the program, the crude birth rate of the population would be reduced from 32 in 1966 to 20 in 1970. Given a death rate of about 5 per thousand, the population growth rates would

([4]) The official criteria for eligibility make no sense when applied to the whole program period because pregnancy and lactation obviously do not endure for five years. We may further note that after setting down these eligibility criteria, the *White Paper* goes on to propose that 15,000 of the first year's target of 25,000 new acceptors will be ex-Family Planning Association clients—the currently practicing contraceptors!—and that efforts should be concentrated in recruiting acceptors from "ante-natal and post-partum mothers"—the currently pregnant or lactating women.

average about 1.5 per cent per year by the end of the planning period. The annual targets of new acceptors were set as follows:

1966	*1967*	*1968*	*1969*	*1970*	*Total*
25,000	30,000	35,000	45,000	45,000	180,000

With the basic infrastructure and groundwork already well developed by the Association over the years, relatively little preparatory work was needed for the inception of the national program. The Board commenced its service by taking over 24 clinics and a number of staff from the Association ([5]). Given an annual fund of $200,000 ([6]) from the Government and the full support of other ministries, the Board quickly expanded its activities. The number of clinics had increased to 33 by the end of the first year and was 49 in 1970.

As originally stated in the *White Paper*, the program was to offer all new acceptors a "menu card" of family planning methods with emphasis on the IUD. During the very first four months, however, a relatively high incidence of uterine translocations was experienced. This, together with exaggerated reports of such side effects as bleeding, cramps and pregnancy, allegedly resulted in a loss of confidence in the method, which threatened to develop into a mistrust of the whole family planning program. Emphasis was therefore shifted to the contraceptive pill, which has remained the most popular method of contraception ever since.

Details on the operation of the national program have been summarized by Chang (1974). Some highlights of its activities include a publicity and educational campaign, a shift of emphasis from "Plan your family" to "Plan your family small", the adoption of "face-to-face" approaches, the development of a post-partum program and the introduction of a family planning advisory service for newlyweds and for industrial and factory workers.

The post-partum program is particularly important because in Singapore 70 to 80 per cent of all births take place in government hospitals and a few more are delivered by the domiciliary delivery service run by the government MCH clinics (see Table 11.10). According to the Singapore Family Planning and Population Board (1970; 2), about 60 per cent of the 17,929 acceptors recruited through the post-partum program in 1970 were completely new cases, while the remaining 40 per cent had been previously registered. The mechanics of this program have been described by Kanagaratnam and Thong

([5]) The Association was left with three clinics to operate until November 1968 when these were also taken over by the Board.

([6]) In Singapore dollars. One US dollar in 1972 was equivalent to approximately 2.8 Singapore dollars.

Kah Leong (1970). The program has had marked success in that nearly 90 per cent of all post-partum and post-abortal women interviewed in hospitals accepted family planning verbally and 40 to 60 per cent actually came to the clinics. Over the five years more than 82,000, or over 50 per cent of all acceptors, were recruited through the post-partum service (Table 11.10), and motivation of post-partum and post-abortal women through government MCH clinics and government hospitals still remains the main source of recruitment of new acceptors.

Table 11.10. *Post-partum project, 1966-70*

Year	Total number of mothers (1)	Total post-partum and post-abortal women in hospitals (2)	Number of women interviewed in hospitals (3)	Number of acceptors (4)	Percentage of pregnancies ending in hospital (2)/(1) (5)	Percentage of women in hospital interviewed about family planning (3)/(2) (6)	Percentage of women interviewed who accepted family planning (4)/(3) (7)
1966	54,310	43,288	40,299	7,679(a)	80	93	
1967	50,158	41,251	39,651	16,512	82	96	42
1968	46,899	39,349	38,256	17,559	84	97	46
1969	44,207	37,480	36,245	22,725	85	97	63
1970	45,596	41,097	37, 793	17,929	91	92	47

(a) Owing to the interruption of the program during the year and its administrative transfer from the joi University-Government Project to the Board, complete data are not available.

Sources: Report on the Registration of Births, Deaths and Marriages, 1966-1970, and SFPPB, *Fifth Annu Report, 1970*, p. 33.

The national family planning program has been supplemented by various policies designed to provide incentives for smaller families and disincentives for large ones. These include the 1968 Employment Act, which stipulates that paid maternity leave is to be given only up to the third confinement, the increase in April 1969 of the accouchement fee charged at government maternity hospitals from $10 to $50 after the third baby, and the granting of income tax relief for dependent children according to a regressive scheme. Moreover, since 1968 couples without any children have become eligible to apply for public housing flats, thus removing an incentive for childbearing from those who may have been planning to have more children, because larger families used to enjoy priority in allotment.

Recently, in order to further discourage parents from having more than two children, the Health Minister announced in Parliament that income tax relief would be reduced from five to only the first three children, that paid maternity leave would be reduced from three to two confinements, and that accouchement charges at government hospitals would be substantially increased. The accouchement fee would be waived, however, if the woman or her husband would undergo sterilization. In addition, larger families would be given low priority in the allotment of public housing flats, thus completely reversing the earlier policy of giving priority to larger families.

There are, in addition, the Voluntary Sterilization Act and the Abortion Act, both of which were introduced in Parliament in late 1968, passed in December 1969, and come into effect in March 1970. The Abortion Act, designed to reform and liberalize the law relating to abortion, provides for legalized abortion not only on purely medical grounds but also on grounds of socio-economic circumstances of the pregnant women. This liberalization, originally recommended by the Family Planning and Population Board, is to allow for abortions as a logical follow-up for contraceptive failures and is not intended as an alternative to contraception. The Voluntary Sterilization Act, on the other hand, serves to clarify the legal position with regard to male and female surgical sterilization. Both Acts are apparently an extension of the earlier family planning legislation and help to boost the national efforts to deal with the problem of excessive population growth.

III. RESPONSE TO THE NATIONAL PROGRAM

An attempt to study the response to the national program was made by taking a 50 per cent sample of the program's new acceptors in the period July 1967 - December 1970. This study has been fully reported by Chang and Yeh (1972) and most of the information in this section is based on their report. Specific studies of the patterns of acceptance of the post-partum IUD program have been done by D. Wolfers (1970b) and H. Wolfers (1970).

The program had an immediate success in the first year when a total of 30,410 new acceptors were registered and nearly one hundred thousand attendances were recorded. The number of new acceptors showed only a very slight increase in the second year, but rose significantly to over 35,000 in 1968. However, it again remained constant in 1969 and following this even dropped considerably to less than 25,000 in 1970, the last year of the program period. In total, 156,556 new acceptors were registered during the five years from 1966 to 1970.

Numbers of New Acceptors, 1966-1970

1966	*1967*	*1968*	*1969*	*1970*	*Total*
30,410	30,935	35,338	35,643	24,230	156,556

Although it had not been able to reach its target of 180,000 new acceptors in the five years, and the crude birth rate of the population, as noted above, had not declined to below 20 per thousand by 1970, the program must nevertheless be considered a success. The annual number of new acceptors averaged more than 31,000, which compares more than favorably with the 9,845 new acceptors registered in 1965, the highest number reached while the family planning program was still under the direction of the Singapore Family Planning Association. The birth rate in 1970 was 22.1 per thousand, which is one-quarter lower than the rate of 29.5 per thousand in 1965 ([7]).

In retrospect, it may be pointed out that the target of 180,000 had been set too high, as it was derived from an over-estimate of married women in the fertile age range of 15-44. The projected female population 15-44, 450,000, compares well with the census count of 455,943 in 1970. However, the nuptiality assumption, which set the proportion of married women in the age group 15-44 at 67 per cent, was obtained without modification from the 1957 census. As observed above, there has since been a strong trend towards later marriage. The proportion was already only 57 per cent in 1966, according to the national sample household survey, and further dropped to 53 per cent by 1970. Thus, because of the erroneous nuptiality assumption, eligible women were substantially over-estimated ([8]).

On the other hand, the target figure of 180,000 was supposed to include only those women who would be currently eligible, that is, who would not be currently practicing contraception or pregnant, lactating, or sterile. As it turned out, more than one-sixth of all new acceptors were actually practicing family planning at the time of their first visit ([9]).

([7]) In attempting to lower the birth rate from 29.5 per thousand in 1965 to 20 per thousand in 1970, the national program in effect aimed at reducing the rate by one-third in five years. It may be noted that in many countries, including South Korea, Taiwan and Ceylon, the goals of their programs have mostly been set as a reduction of the same proportion in ten years. See L. P. Chow and S. C. Hsu (1969).

([8]) Applying the eligibility assumption of 60 per cent to the census count of married women 15-44, there would be only about 144,000 eligible women. The number of 180,000 is then an overestimate by 25 per cent.

([9]) See footnote 4.

GENERAL CHARACTERISTICS OF ACCEPTORS

C. T. Chang (1974) has described the ethnic group, age, education, economic activity, occupation, and income of acceptors and some characteristics of their husbands. Although space limitations prohibit more than a short summary of these data, they help to illuminate the way in which the national program has affected reproductive behavior. Suffice it to say that over the five-year period, the ethnic composition of new acceptors has become representative of the total population; their age has declined; their educational attainment has, in general, been representative, with an increase in educational attainment over the period; their labor force participation has risen, although it has remained below the general level of economic activity; the occupation of their husbands, except for service workers, has been reasonably representative and their family income has increased. It is clear that the program has reached all major sections and strata of the population. If there were "hard core" resisters to family planning they were not concentrated in any readily measurable socio-economic segments of the population. Indeed, a recent ECAFE study in Singapore (Chen, 1972) reveals that socio-economic distribution of the "hard core" group, i.e., couples who have never used any method of family planning, hardly differs from that of the overall national sample.

CONTRACEPTIVE BACKGROUND

Not all the SFPPB new acceptors were new to the practice of contraception, and some categories were more likely than others to have had contraceptive experience. Of all the new acceptors registered between 1967 and 1970, 71 per cent had never practiced contraception before. Those who were in fact using contraceptives at the time of their first visit to a SFPPB clinic constituted 18 per cent of all new acceptors, while 10 per cent had previously used contraceptives but for some reason had discontinued use (Table 11.11).

Compared with the Malay and the Indian and Pakistani new acceptors, the Chinese were more likely to have had previous contraceptive experience on their first visit to the SFPPB clinic.

PARITY

One outstanding feature of the new acceptors is the large percentage having only a small number of children at the time of their first visit (Table 11.12). This, of course, is largely attributable to the high proportion of spacing among contraceptors. Nearly one

Table 11.11. *Percentage without contraceptive experience by ethnic group and number of living children, 1967-70*

	Number of living children				
	0-1	*2*	*3*	*4*	*5 or more*
All races	87	69	62	61	62
Chinese	90	70	60	57	55
Malays	93	87	82	79	80
Indians and Pakistanis	91	75	74	75	67

Source: C. T. Chang and Stephen Yeh, 1972.

Table 11.12. *Percentage distribution of new acceptors by number of living children and selected characteristics, 1967-70*

	Number of living children				
	0-1	*2*	*3*	*4*	*5 or more*
All races	30	19	14	10	27
Ethnic group					
Chinese	33	18	13	10	26
Malays	19	15	12	11	43
Indians and Pakistanis	23	19	17	13	28
Education					
No education	13	13	13	13	48
Primary	30	21	15	12	22
Secondary or higher	53	27	11	5	4
Combined income (a)					
Less than S$200	21	16	14	12	37
S$200-S$400	29	20	13	11	27
S$400-S$600	41	22	14	8	15
More than S$600	46	30	14	5	5

(a) Figures in Singapore dollars.

Source: C. T. Chang and Stephen Yeh, 1972.

half of the acceptors registered between 1967 and 1970 were women with *at most* two children, and 30 per cent had one child or none at all. These figures compare favorably with what has been found in several Asian countries where there have been family planning programs for some time.([10])

Parity composition differs significantly according to socio-economic status largely in the expected manner (Table 11.12). Of the three main ethnic groups, the Chinese have the highest proportion of acceptors with only one child or none at all, followed by the Indians and Pakistanis; at the other end of the scale the Malays have the highest proportion (43 per cent) of acceptors with at least five children at the time of their first visit to the clinics.

REASON FOR FAMILY PLANNING

A new acceptor may want to stop having children altogether, or she may wish only to postpone further births. In the first case, the reason for family planning will be *limiting*; in the second, *spacing.* Of course, it is also possible that, at the time of her first visit, she might not have made up her mind whether to space or to limit her childbearing. In such cases, the answer to the question on the reason for family planning will be "undecided".

Since spacers have less than the number of children that they want and limiters have at least as many as they want, a difference in mean parity is to be expected between the two. In fact, some 50 per cent of the new acceptors reporting spacing as their purpose for family planning had only one child or none at all, and fully three quarters had two children or fewer. On the other hand, more than 60 per cent of those who want to stop having children completely had at least five children. Of all the new acceptors registered between 1967 and 1970 nearly 60 per cent came before completing their family-building while those reporting limiting as their purpose account for 38 per cent. Child-spacing was thus the main reason for family planning.

Just as the current family size of new acceptors differs according to their socio-economic characteristics, so does their composition with regard to spacing or limiting future births. Indeed, much of the difference in the number of children of the various socio-economic groups of acceptors can be explained by their differences in reason for family planning.

([10]) In Taiwan, for example, it was found in 1968 that women with two children or less made up less than 20 per cent of all new loop and pill acceptors for that year. The comparison is even more striking if we look at the proportion with one or no child. While this proportion is 30 per cent in Singapore, the figure for Taiwan is only 4 per cent.

DESIRED FAMILY SIZE

Not only do the different socio-economic groups among new acceptors vary in the timing of their family planning but they also differ in the completed family size they consider desirable. Although direct information is not available on this variable, it is possible from the data to derive some idea of the relative desired family sizes of different groups of new acceptors. Given the same number of living children at the time of their first visit, women who come for family planning services with a larger desired family size in mind are naturally more likely to report spacing as their purpose than are those whose desired family size is smaller. On the basis of this assumption, Chang and Yeh (1972) compared the distribution by reason for family planning of the different groups of new acceptors, while holding constant the number of their living children. Those groups of acceptors showing a higher proportion wishing to limit births are taken as having a smaller desired family size; correspondingly, those with a higher proportion wishing to space are interpreted as desiring larger families.

Of course, the average desired family size of a group of acceptors is determined not only by that of the limiters but also by the views of the spacers in the same group. This means that, in making inferences from the difference in the proportion wishing to limit births to the relative desired family sizes of the different groups, it must be assumed that the desired size does not differ between the spacers of any two groups in such a way that these inferences would be invalidated. Because of this plausible but unverified assumption, the findings must be interpreted carefully and perhaps tentatively. Only proportional differences of some magnitude may be taken as evidence of a difference in the desired family size between groups of new acceptors.

The analysis made showed that the Malays desired a larger number of children than the Chinese or the Indians and Pakistanis (see Table 11.13), that the desired number of children, in general, varied inversely with various indicators of socio-economic status and that the sex preference for children was such that women were more likely to be limiting if they had children of both sexes or if they had one, or especially two, sons (Table 11.14).

A comparison over time further revealed that spacing was becoming a more and more common reason for acceptance of family planning and that, as a consequence, the parity of acceptors was shifting downwards; that the desired number of children was declining, especially among women of the postwar baby boom cohorts in their reproductive ages; and that there was a decrease in the proportion of acceptors without previous contraceptive experience.

Table 11.13. *Percentage of new acceptors limiting by number of living children and selected characteristics, 1967-70*

	Number of living children			
	0-1	*2*	*3*	*4*
All races	1	15	38	66
Ethnic group				
Chinese	1	13	38	70
Malays	1	5	20	46
Indians and Pakistanis	1	15	41	69
Education				
No education	2	13	37	67
Primary	1	10	36	66
Secondary or higher	1	21	52	75
Combined income (a)				
Less than S$200	1	10	33	64
S$200-S$400	1	13	41	71
S$400-S$600	2	19	54	77
More than S$600	2	29	65	89

(a) Figures in Singapore dollars.
Source: C. T. Chang and Stephen Yeh, 1972.

Table 11.14. *Percentage of new acceptors limiting by number of living children and number of living sons, 1967-70*

Number of living sons	*Number of living children*				
	0-1	*2*	*3*	*4*	*5 or more*
0	1	8	13	26	39
1	1	19	35	58	71
2		12	48	75	86
3			34	73	87
4 or more				49	90

Source: C. T. Chang and Stephen Yeh, 1972.

The success of the national program is attributable in some part to the compact physical setting of life in Singapore where the rural-urban distinction has verged on irrelevance, to the social improvement and rapid economic development in recent years, to the consequent relatively high standards of living and, not least important, to the preparatory work done by the Singapore Family Planning Association in building up receptivity to the concept and practice of family planning among the general public.

IV. SELECTED FEATURES OF THE PROGRAM

Also described in the work of C. T. Chang (1974) are various significant features of the national program experience. Space does not permit elaboration here, but a brief summary of the findings is presented.

ACCEPTANCE RATIOS

As the first five-year period of the national family planning program drew to its end in 1970, 58 per cent of married women, or about one third of all women within the reproductive age-range, had taken part in the program (Table 11.15). The percentage of acceptors has been especially high for married women in their twenties, of whom four out of five have attended the clinics and requested family planning services. Since in recent years the average age of women at first marriage has been about 23, the high acceptance ratio indicates that young couples are fully aware of the importance of family planning and many have approached the clinics very soon after their marriage. It will be noted that, even among married women under 20, one in two has sought family planning services from the program. The acceptance ratio decreases with increasing age for married women 30 years of age and above. It is nearly 70 per cent for married women aged 30-34 but only 14 per cent for those aged 45-49.

The percentage of all women who have been acceptors in the program hardly differs among the three ethnic groups. It is 34 per cent for the Malays and 31 per cent for both the Chinese and the Indians and Pakistanis. The acceptance ratio of married women varies, however. Whereas 59 per cent of married Chinese women have taken part in the program, only 50 per cent of the Indian and Pakistani married women have ever participated. The ratio for the Malays lies in between at 54 per cent.

Age-specific comparisons reveal that the acceptance ratio of married women under 20 is significantly higher for the Chinese than

Table 11.15. *Ratio of estimated surviving new acceptors to currently married women in 1970, by age and ethnic group* [a]

Age group in 1970	Total population		Chinese		Malays		Indians and Pakistanis	
	Surviving new acceptors	% of married women	Surviving new acceptors	% of married women	Surviving new acceptors	% of married women	Surviving new acceptors	% of married women
15-19	2,911	50.7	1,905	59.3	699	37.6	256	45.6
20-24	28,417	80.8	19,575	81.3	5,299	72.6	2,127	73.9
25-29	41,392	82.2	31,609	82.2	5,718	75.9	2,612	82.6
30-34	40,960	68.7	32,608	72.7	5,497	57.5	1,957	51.6
35-39	24,971	51.1	19,454	52.3	4,036	53.5	1,041	35.1
40-44	13,130	32.2	10,564	33.5	1,932	32.3	468	20.3
45-49	4,375	14.3	3,559	14.7	601	14.6	110	7.4
15-49	156,156	57.6	119,274	58.6	23,782	54.2	8,571	50.0

(*a*) The age distribution of new acceptors in 1966 is assumed to be the same as that of new acceptors in 1967 (to be exact, July-December 1967). The estimates of surviving new acceptors in 1970 are obtained by applying appropriate survival ratios derived from 1970 life tables to new acceptors of each year.

for the other two groups, and that married Indian and Pakistani women aged 35 and over have markedly lower percentages accepting than their Malay and Chinese counterparts.

It should be emphasized that the differences in acceptance ratios must be interpreted carefully. In a predominantly post-partum program, the expected acceptance ratios are affected by both fertility levels and the proportion of deliveries taking place in hospitals. In Singapore the proportion of deliveries in hospitals is highest among the Chinese and lowest among the Malays. On the other hand, Malay fertility has been the highest and that of the Chinese the lowest among the three major ethnic groups. Thus, these two factors tend to an unknown extent to cancel each other out in their effects on the acceptance ratio.

NON-PROGRAM ACCEPTORS, ABORTIONS AND STERILIZATION

In the five years reviewed a large number of women have reportedly attended clinics other than those of the national program. As mentioned previously, the Association retained three clinics when the Board took over the responsibility for the family planning program in 1966. These three clinics were maintained by the Association until November 1968, when they were also transferred to the Board. During this period of nearly three years, a total of 4,511 new acceptors had been recruited by the Association. In addition to the three clinics, the regional headquarters of the International Planned Parenthood Federation (IPPF) in Singapore had operated two fixed clinics in this period and one mobile clinic until May 1970. From 1966 to May 1970 nearly three thousand women had been registered as new acceptors at these clinics. There are also, of course, many women who prefer to see private doctors for contraceptive advice and services. Their exact number is, however, difficult to ascertain, though in the first three years the Board took care to obtain reports from private medical practitioners on their provision of family planning services. On the basis of these far from complete reports, it has been recorded that more than 24,000 women obtained contraceptives from this source during the three years. The number of new acceptors so recorded cannot, however, be readily added to the number of new acceptors of the national program. This is because the practice of the Board, and for that matter of the other two agencies and private doctors as well, is to record a woman as a new acceptor when she first attends the clinics, regardless of her contraceptive background. Moreover, after the Board took over the remaining three clinics of the Association in November 1968, all the Association's existing cases who revisited the clinics were defined as the Board's new acceptors.

It is significant to observe that the number of such non-program acceptors has declined since 1966. Although there are problems of administrative transfer or data collection involved, the downward trend may also suggest that, as the national program became established, it started to recruit more and more women who would otherwise have approached the other sources for family planning services.

Apart from the acceptance of contraception, a number of abortions and sterilizations have also been performed in recent years. Abortion had previously been quite common in Singapore. Long before the introduction of the Abortion Act, when the family planning program was under the direction of the Association, it had been noted that many women came to the clinics to ask for abortions and a fair proportion of new acceptors admitted that they had had one or more abortions before attending the clinics. There are no official figures of the total number of illegal abortions in Singapore, and even figures on abortions performed in government hospitals are available only for recent years. These figures are presented in Table 11.16. It can be seen from the table that, until 1970, there was a downward trend in the number of abortions performed in government hospitals. Since the downward trend occurred in the period after the national program was implemented, it would seem that one of the results of the program was to reduce the number of unwanted conceptions which would have resulted in abortions. The substantial increase in the number of abortions in 1970 must undoubtedly be attributed to the implementation of the Abortion Act. It remains to be seen whether the liberalization of the abortion law will have adverse effects on the practice of contraception, notwithstanding the government's claim that abortion is not to be a substitute for family planning.

Table 11.16. *Number of female sterilizations and abortions, 1964-70* ([a])

Year	*Female sterilization*	*Abortion*
1964	557	4,295
1965	542	3,386
1966	477	3,432
1967	653	3,327
1968	1,054	2,922
1969	1,435	2,929
1970	2,310	5,300 ([b])

([a]) Sterilization and abortion performed at Kandang Kerbau Maternity Hospital and Thomson Road General Hospital.

([b]) Includes 1,886 abortions legalized under The Abortion Act, 1969.

Source: SFPPB, *Fifth Annual Report, 1970*, and S. B. Kwa, S. T. Quah and Mark C. E. Cheng, "The Abortion Act, 1969—A Review of the First Year's Experience," Singapore Medical Journal, 12: 5 (October 1971).

As may be expected, it has been found that there has been a definite bias in the actual implementation of the Abortion Act in that it has been more difficult to obtain abortions on socio-economic grounds for women under 30 years, of low parity, or in high income groups (Kwa *et al.*, 1971).

In addition to abortions, a number of women have also been sterilized in recent years. The number, moreover, has been increasing over the program period. Whereas in 1966 the number was only 477, it rose to 1,435 in 1969 and further to 2,310 in 1970 when the Voluntary Sterilization Act was implemented. For the five years from 1966 to 1970, a total of 5,929 women underwent sterilization, compared with 17,910 abortions during the same period. Male sterilization has been unpopular, however. There were only three cases in 1968 and three in 1969. In 1970 the number increased to about 40.

Kwa *et al.* (1971) also found that a fair proportion, 30 per cent, of the abortions approved and performed during the first ten months of the implementation of the Abortion Act was associated with a simultaneous sterilization operation.

CONTINUATION OF FAMILY PLANNING

Attempts have been made to study the continued use of contraceptive methods acquired through the national program (Wan Fook Kee and Quah Siam Tee, 1971; Kanagaratnam and Khoo Chian Kim, 1969; D. Wolfers, 1969). According to these studies, 54 per cent of pill acceptors are still obtaining pills from the clinics after an average of 12 months and 39 per cent after 24 months. As expected, the status of program use is related to the number of living children. Whereas only 24 per cent of oral contraceptors with one or no children are still on the pill after two years, the proportion is 40 per cent for those with two children and 43 for those with three or more. An obvious explanation is, of course, that women with only one child or none at all are much more likely spacers than are those with more children, as has been observed above. The same factor must also be involved in the ethnic differences in program use of pill acceptors —38 per cent continuation for the Chinese, 48 per cent for the Malays, and 23 per cent for the Indians and Pakistanis, after an average of two years—but the preference of the Malays for oral pills may also be relevant here. Little difference exists in this respect between women without any formal education and those with at least some education, and the slightly lower percentage of the latter's continuing program use may be explained by their higher proportion of those wishing to space.

The main causes of discontinuation of oral contraceptive use are medical reasons and pregnancy, planned and unplanned, with medical reasons accounting for 25 per cent of all program dropouts or about 40 per cent of those for whom reason for termination is known. Among all medical reasons given, nausea, vomiting, and giddiness are the most common, and none of the dropouts has been by reason of thrombosis. Pregnancy is the reason given by 9 per cent of all dropouts or 15 per cent of those with known reason for termination. Of the pregnancies occurring, 83 per cent have been found to be planned.

Of course, not all pill acceptors who have discontinued obtaining supplies from the clinics have abandoned contraception. Of the 60 per cent discontinuers after 24 months for whom information is available, 44 per cent have stopped practicing contraception for reasons including pregnancy, while 56 per cent are using alternative methods of protection. In most cases the alternative method is the condom.

Using a technique derived from Tietze's method, Wan Fook Kee and Quah Siam Tee (1971) estimated that the cumulative continuation rates of pill acceptors are 689, 500, and 436 per thousand women at the end of the third, twelfth, and eighteenth cycles respectively. After the twenty-fourth and thirtieth cycles, the rates are 372 and 332.

Wolfers (1969), in his study of the demographic effects of the post-partum IUD project, finds that, after a mean period of 16.44 months, the proportion continuing IUD use is 64 per cent for the Chinese, 84 per cent for the Malays, and 76 per cent for the Indians and Pakistanis. The ethnic differences persist for practically all age groups. Also applying Tietze's method, he further estimates that the cumulative continuation rate is 867 per thousand women at the third month, 688 at the twelfth month, and 652 at the eighteenth month.

V. THE NATIONAL PROGRAM IN RELATION TO FERTILITY CHANGES

To relate the national program to fertility changes in recent years, it must be noted that the decline in the total fertility rate of the population has accelerated substantially since 1965. Whereas between 1957 and 1960 the total fertility rate decreased by 3.1 per cent per year and between 1960 and 1965, by 3.7 per cent per year, the rate of reduction nearly doubled to 7.0 per cent per year from 1965 to 1970. This acceleration of fertility decline was especially conspicuous in

Table 11.17. *Changes in total fertility rates and the net effects of component changes by ethnic group and period*

Ethnic group and component	*Amount of change*			*Percentage of change*		
	1957-60	*1960-65*	*1965-70*	*1957-60*	*1960-65*	*1965-70*
Total population	−596	−1,083	−1,682	−9.2	−18.5	−35.1
Nuptiality	−325	−556	−379	−5.0	−9.5	−7.9
Marital fertility	−264	−480	−1,358	−4.1	−8.2	−28.4
Chinese	−841	−1,371	−1,300	−12.8	−24.0	−30.0
Nuptiality	−319	−517	−295	−4.9	−9.1	−6.8
Marital fertility	−531	−827	−1,030	−8.1	−14.5	−23.8
Malays	321	−337	−2,759	5.1	−5.1	−44.1
Nuptiality	−348	−668	−755	−5.5	−10.1	−12.1
Marital fertility	724	396	−2,199	11.5	6.0	−35.1
Indians and Pakistanis	197	−941	−3,301	2.7	−12.7	−50.8
Nuptiality	−422	−796	−982	−5.8	−10.7	−15.1
Marital fertility	675	−138	−2,573	9.3	−1.9	−39.6

Source: Chang (1974).

Table 11.18. *Changes in crude birth rates and the net effects of component changes by ethnic group and period*

Ethnic group and component	*Amount of change*			*Percentage of change*		
	1957-60	*1960-65*	*1965-70*	*1957-60*	*1960-65*	*1965-70*
Total population	−5.2	−8.0	−7.4	−12.2	−21.3	−25.1
Age composition	−1.4	−0.4	4.3	−3.3	−1.1	14.6
Nuptiality	−2.6	−4.4	−2.2	−6.1	−11.7	−7.5
Marital fertility	−1.2	−2.4	−7.3	−2.8	−6.4	−24.7
Chinese	−6.5	−8.5	−5.6	−15.3	−23.5	−20.3
Age composition	−0.8	0.1	3.1	−1.9	0.3	11.2
Nuptiality	−2.7	−3.9	−2.3	−6.3	−10.8	−8.3
Marital fertility	−3.2	−4.1	−5.7	−7.5	−11.4	−20.7
Malays	−2.4	−6.7	−16.7	−4.8	−14.1	−41.0
Age composition	−4.4	−3.5	3.6	−8.8	−7.4	8.8
Nuptiality	−3.7	−5.4	−5.4	−7.4	−11.4	−13.3
Marital fertility	5.7	2.9	−13.9	−11.4	6.1	−37.0
Indians and Pakistanis	−0.1	−6.0	−10.0	−0.3	−16.8	−33.6
Age composition	−1.3	−0.5	12.8	−3.6	−1.4	43.0
Nuptiality	−2.7	−4.1	−5.0	−7.5	−11.5	−16.8
Marital fertility	3.8	−0.1	−11.0	10.6	−0.3	−37.0

Source: Chang (1974).

the two ethnic minorities. Among the Malays, the reduction increased from 1 per cent per year in the period 1960-65 to 8.8 per cent per year in 1965-70. For the Indians and Pakistanis, the acceleration went from 2.5 per cent per year in 1960-65 to as much as 10.2 per cent per year during the program period. It seems unlikely that such a magnitude of change in the pace of fertility decline could have occurred without the implementation of the national family planning program and the related government policies.

The role of the national program in the recent acceleration of fertility transition will be better understood after examining the sources of changes in fertility. For this purpose a decomposition procedure based on the component framework expounded by Kitagawa (1955) has been employed which permits an analysis of the changes in total fertility rates or crude birth rates in a given period as effected by changes (a) in the marital pattern of women of fertile ages, (b) in the fertility of married women, and (c) for the crude birth rate in the age and sex composition of the population during the period. Tables 11.17 and 11.18 present the findings from such component analyses of the changing total fertility rates and crude birth rates respectively, for the total population as well as for the three ethnic groups and for three time intervals between 1957 and 1970.

It will be observed from Table 11.17 that whereas in both the 1957-60 and the 1960-65 periods, more than half of the declines in total fertility rates resulted from changes in the marriage pattern of women in fertile ages, during the period 1965-70 as much as 80 per cent of the decline was due to changes in fertility within marriage. Indeed, the acceleration in the reduction of the total fertility rate would not have occurred if the decline in marital fertility had not accelerated substantially after 1965. The effect of the changing marriage pattern on the total fertility rate of the population was a reduction of an average 1.7 per cent per year in 1957-60, 1.9 per cent per year in 1960-65, and 1.6 per cent per year in 1965-70. In contrast, the net effect of changing marital fertility increased from a reduction in the total fertility rate by an average of 1.4 per cent per year in 1957-60 and 1.6 per cent per year in 1960-65 to a reduction of 5.7 per cent per year in 1965-70.

The significant impact of the substantial reduction of marital fertility during the program period stands out even more clearly if the changing age composition is considered in relation to changes in the crude birth rate. As Table 11.18 shows, the crude birth rate of the population declined by an average of 4.1 per cent per year during 1957-60 and 4.3 per cent per year during 1960-65. For both periods the main source of the reduction was the changes in nuptiality of the

female population, but both the changes in age composition and in marital fertility also operated to reduce the birth rate. From 1965 to 1970, however, whereas the changing marital pattern alone could have reduced the birth rate by 7.5 per cent, the changes in age composition, owing to the entry of the postwar boom babies into their twenties, could on the other hand have pushed the birth rate up by as much as 14.6 per cent. It is the large decline in fertility within marriage during this period, which by itself operated to reduce the birth rate by 24.7 per cent, that more than offset the effect of changing age composition and which, together with the negative impact of nuptiality, brought the birth rate down by 25.1 per cent in these five years.

Fertility changes and their major sources differed before 1965 among the three ethnic groups but especially between the Chinese and the two ethnic minorities. All three groups, however, shared the same experience of a substantial acceleration in marital fertility reduction after 1965. For all three groups, also, the reduction of fertility within marriage was the predominant source of their fertility decline after the commencement of the national family planning program.

Although it is difficult to make a precise assessment of the extent to which the program contributed to the accelerated decline in fertility, that an acceleration on this scale should have occurred in the period when the national program was implemented is certainly more than coincidental. And if the accelerated reduction in marital fertility is indeed genuine, as it appears to be, it would be very difficult to deny that there must be some causal connection between this acceleration and the implementation program. To be sure, the reduction in marital fertility increased between 1957-60 and 1960-65, as it did in total fertility, but the pace of acceleration in the earlier periods is minor compared with that which occurred after 1965.

Further evidence in support of the causal relation between the program and the recent fertility changes is apparent in the great increase in the decline beginning in 1967 (Table 11.19). Although the program started in January 1966, the total fertility rate in 1966 was only 5.9 per cent lower than in 1965. This decrease is of about the same order of magnitude as the reductions of 3.9 per cent in 1963-64 and 4.2 per cent in 1964-65. It may be argued that the 1966 decrease could well have been expected from simple extrapolation of past fertility changes even without the national program. From 1966 to 1967, however, total fertility declined by 12 per cent, and was followed by a further reduction of 9.9 per cent in 1967-68. Since a family planning program cannot produce an impact on births in less

than nine months from its commencement and is likely to show significant effects on fertility only after the first year, it is plausible to assume that the relatively slow rate of decline in 1966 reflected precisely this first-year "moratorium", and to interpret the dramatic decline in fertility in 1967 and 1968 as the result of the massive increase in the number of new acceptors in the first two years of the national program. This conclusion is supported by D. Wolfers (1970a) who also attempted to evaluate the impact on fertility of the national program by extrapolating the fertility trend in 1961-1965. He found little difference between the actual and the projected births for 1966, and significant differences emerged only after the first year of the program.

Table 11.19. *The total fertility rate, 1957-70*

Year	*Total fertility rate*	*Percentage of decline*
1957	6,465	
1958	6,422	0.7
1959	6,230	3.0
1960	5,868	5.8
1961	5,506	6.2
1962	5,258	4.5
1963	5,199	1.1
1964	4,994	3.9
1965	4,786	4.2
1966	4,506	5.9
1967	3,965	12.0
1968	3,574	9.9
1969	3,257	8.9
1970	3,104	4.7

Source: Chang (1974).

It is of interest in this context to report another study by D. Wolfers (1969), this time in assaying the impact of the joint government-university post-partum program at Kandang Kerbau Hospital. The program was conducted from October 1965 to July 1966. During this period some 9,000 women accepted the IUD. After a hiatus of four months the program was resumed in November 1966 with a multiple choice of contraceptives. In his careful analysis of the time-series data on the monthly registrations of births, Wolfers noticed that

there were two sharp plunges in registered births, one in January 1967 and the other in February 1968. Since the results of a post-partum program will not become apparent before fifteen months after its commencement, Wolfers maintains that the occurrence of these two large dips in registered births exactly fifteen months after the two commencements of post-partum activity suggests a more than coincidental relationship between the program and the decreases in live births.

As would be expected, the accelerated decline in fertility in 1965-70 was associated with a major reduction at higher orders of birth. From 1967 to 1970, as Table 11.20 shows, while fertility for the first three orders of birth declined on the average by about 3 per cent per year, fertility for higher-order births, especially those of the fifth order and above, decreased by 15 per cent or more per year. Indeed, the decreases in births of the fourth order and above would in themselves have effected a reduction of about 18 per cent in the total fertility rate. Thus, more than four-fifths of the total fertility decline of 22 per cent in the four years must be attributed to the decrease in fourth and higher-order births.

As a result of the differential changes in fertility by birth order, the pattern of fertility of the population by birth order has changed.

Table 11.20. *Gross total fertility by birth order, 1967-70* (a)

Birth order	*1967*	*1968*	*1969*	*1970*	*Percentage of change*		
					1967-1968	*1968-1969*	*1969-1970*
1	162.4	156.7	152.7	148.0	−3.5	−2.6	−3.1
2	141.0	136.3	135.0	132.2	−3.3	−1.0	−2.1
3	105.6	101.3	100.5	98.3	−4.1	−0.8	−2.2
4	84.9	73.2	67.0	65.9	−13.8	−8.5	−1.6
5	69.4	59.5	49.2	45.1	−14.3	−17.3	−8.3
6	58.2	49.1	37.7	34.1	−15.6	−23.2	−9.5
7	49.9	39.5	31.5	28.5	−20.8	−20.3	−9.5
8	38.0	32.4	25.2	22.0	−14.7	−22.2	−12.7
9	27.7	23.0	18.1	15.7	−17.0	−21.3	−13.3
10+	49.2	39.9	31.5	26.5	−18.9	−21.1	−15.9

(a) The gross total fertility rate for live births of a given birth order is simply the sum of the age-specific fertility rates for births of that order computed by five-year age interval. The sum of the order-specific GTF's should but does not equal one-fifth of the TFR because, in computing these order-specific GTF's, both births of unknown birth order and births of unknown age of mother are excluded. In computing age-specific birth rates and hence TFR, births with unknown age of mother have been allocated proportionally according to the distribution of births by age of mother.

Source: Chang (1974).

Whereas in 1967 the first three orders of live births accounted for 52 per cent of the total fertility of the population, by 1970 their relative contribution had increased to 62 per cent (Table 11.21). This pattern of changes in order-specific fertility is characteristic of a population in which couples with many children are making an effort to limit the size of their families.

Of course, not all the reduction in higher-order births has resulted from family limitation; part of it is attributable to the postponement of childbearing. It will be noticed that whereas fertility has been declining for every order of live births in 1967-70, major reductions have increasingly shifted to higher orders of births (Table 11.20). For instance, while fertility for fourth births showed a reduction by roughly 14 per cent in 1967-68, the decrease was reduced to 8.5 per cent, and further to only 1.6 per cent in the following two years. Similarly, although the fertility rate for births of the fifth order declined by 14.3 and 17.3 per cent in 1967-68 and 1968-69 respectively, in 1969-70 the reduction was lowered to only 8.3 per cent. Such an upward shift by birth order of the major reductions in fertility is indicative of a process whereby early reductions in fertility, owing to the practice of child-spacing, are somehow "compensated" for later as the delayed births appear.

The significance of the wide practice of child-spacing that has been observed above must be assessed along the tempo as well as

Table 11.21. *Relative contribution of fertility for each order of live births to the total fertility rate, 1967-70*

Birth order	*1967*	*1968*	*1969*	*1970*
1	20.7	22.0	23.6	24.0
2	17.9	19.2	20.8	21.5
3	13.4	14.2	15.5	16.0
4	10.8	10.3	10.3	10.7
5	8.8	8.4	7.6	7.3
6	7.4	6.9	5.8	5.5
7	6.3	5.6	4.9	4.6
8	4.8	4.6	3.9	3.6
9	3.5	3.2	2.8	2.5
10+	6.3	5.6	4.9	4.3
All	99.9 (a)	100.0	100.1 (a)	100.0

(a) Due to rounding.
Source: Chang (1974).

the quantity dimension of fertility. Child-spacing, by definition, implies merely a temporal redistribution of live births over a period of several years; its impact on fertility, as far as the quantity dimension is concerned, is therefore of only transitory a nature. To be sure, some births delayed because of spacing may be permanently averted owing to the death or separation of the potential parents or the onset of menopause of the potential mothers. Such "births", however, are not likely to be more than a small fraction of the total births postponed. Because, in effect, relatively few births will thus be prevented and most of the delayed births will come later anyway, the period measures of the quantity of fertility may simply fluctuate from year to year without showing any definite downward tendency. Indeed, the upsurge in births in 1971 and 1972 after great decreases in 1966 to 1969 must in part be attributed to previously postponed births during the first three or four years of the national program, as well as to the sloughing off of new acceptors after the initial program impact.

As a matter of fact, the reduction in the total fertility rate has decelerated persistently following the plunge of 12 per cent in 1967, to 9.9 per cent in 1968 and 8.9 per cent in 1969. From 1969 to 1970, the decrease in the total fertility rate further fell to a decline of only 4.7 per cent (Table 11.19).

To appreciate better the significance of the practice of child-spacing, the tempo dimension of fertility must be considered in long-term perspective. The first observation to be made in this respect is that the spacing of births implies a shift upwards in the distribution of births by age of mother. The mean age of childbearing being its principal determinant, the length of a generation is prolonged in proportion to the average number of years for which births are delayed because of child-spacing. Since the intrinsic rate of natural increase varies inversely with generation length, the long-range effect of child-spacing is to reduce the rate of population growth. The point has been simply phrased by Ryder (1957) who observed that "a birth to a woman at age 20 is worth approximately twice as much as a contribution to replacement as a birth at age 40".

Table 11.22 presents the mean age of fertility, defined as the mean of the distribution of fertility rates by age, for each order of live births for the period 1967-70. The mean age of fertility has been computed specifically for each order of birth because, as has been noted, high orders of births have been declining much faster recently than lower orders, and such a pattern of change has the effect of reducing the mean age of fertility for all births independently of any change in the tempo dimension of fertility. It can be seen from the table that for births of all orders the mean age of fertility has indeed

been rising continuously since 1967, in some cases by as much as a whole year. Although rising age at marriage must have contributed to this development, such a large change in fertility tempo within such a short period is not likely without increased child-spacing.

Table 11.22. *Mean age of fertility by birth order, 1967-70*

Birth order	*1967*	*1968*	*1969*	*1970*	*Increase (1967-70)*
1	24.3	24.6	24.8	25.0	0.7
2	25.8	26.3	26.5	26.8	1.0
3	27.4	27.7	28.1	28.5	1.1
4	28.9	29.1	29.5	29.8	0.9
5	30.5	30.8	30.8	31.2	0.7
6	32.1	32.4	32.6	32.9	0.8
7	33.8	33.8	34.0	34.4	0.6
8	35.3	35.4	35.7	35.8	0.5
9	36.5	36.7	37.0	37.3	0.8
10+	38.7	38.8	38.9	39.1	0.4

Source: Chang (1974).

VI. DISCUSSION AND CONCLUSION

The decline of fertility in Singapore is often heralded as one of the relatively few dramatic successes of a national family planning program in developing countries. The evidence presented above certainly confirms the conclusion that the family planning program in Singapore did effectively contribute to a sharp drop in the birth rate, especially between 1965 and 1970. But the evidence also indicates that other factors played a significant role. Among these factors is the relatively rapid modernization experienced by this city-state since the end of the Second World War, with concomitants that, in addition to the family planning program, contributed to the decline in the birth rate. These included both socio-economic and demographic factors contributing to lower fertility. The former was evident in increased education and higher levels of living; the latter in decreased nuptiality by reason of the rising age at marriage and in the changing age structure. Furthermore, there is clear evidence that the decline in Singapore's birth rate had begun before the Family Planning Association or the government program had gained sufficient momentum to have significant impact. The initiation of the decline in fertility in

Singapore cannot, therefore, be attributed to the family planning program.

Moreover, since the birth rate began to fall before family planning clinics reached significant proportions of women, it is in order to examine the question of whether the birth rate would have continued its downward course and reached its present level without the program. That is, whether the social and economic changes that among other things generated the program would perhaps have led acceptors to restrict family size even if the program had not been available. If this is true, then the program would have served only as a substitute for non-program sources of birth control and the observed fertility effects would have occurred even without it. It has been suggested, however (Freedman *et al.*, 1972), that such substitution is more relevant in modern low fertility areas and less relevant in traditional high fertility areas, because in the former the greater demand for fertility control methods has probably resulted in their greater availability from various sources.

It will be recalled that as many as 30 per cent of the national program's new acceptors were currently practicing or had discontinued usage of birth control methods. Some of the latter undoubtedly gained their family planning experience through the clinics run by the Family Planning Association before the national program was launched. Some of them, on the other hand, could have obtained family planning services from other sources, especially from private medical practitioners. It has also been noted that new acceptors who attended IPPF and SFPA clinics, as well as those recruited by private medical practitioners, have all decreased in number since the inception of the national program. The number probably declined because more and more women who could have obtained family planning services from these sources attended the Board's clinics instead, as the national program became more widely known. This is the more likely in view of the Board's policy of keeping charges comparatively low for contraceptive materials and services. Thus, although it is difficult to determine the extent to which the national program has served only as a substitute for other possible sources, there is good reason to believe that some such substitution occurred in the Singapore experience.

To refer to such possible substitution effects, however, is not to offset the evidence that the national program did make a major contribution to the reduction of fertility.

For one thing, the publicity campaigns of the Board through mass media and personal contacts contributed much to the dissemination of family planning information. It has been found in a recent study

(Chen, 1972) that hardly more than one in ten married people has not heard about family planning. Many persons first came to know about family planning from health and family planning workers rather than from other sources. Moreover, for many people and especially for the lower socio-economic groups, the program workers were often their most important personal, as distinct from mass media, source of information. Further, the Board has extensively used the government radio and television network for its publicity campaigns. Newspapers also gave repeated and broad coverage to the program and related family planning information and it is doubtful that all this publicity would have been possible in its absence.

Secondly, while the lower charges for material and services at the Board's clinics could have accounted for part of the substitution effects of the national program, its low charge policy undoubtedly helped to extend services to the high-fertility families in the lower socio-economic strata of the population. It is significant that women of lower socio-economic status were not under-represented among the program's acceptors.

There is, moreover, a component of the national program which has proved of especially great value—the post-partum motivation project instituted in government maternity hospitals. The project seeks to explore and tap the special advantages in accessibility and motivation of women afforded by the post-partum situation. It has been especially important in Singapore because from three-quarters to four-fifths of births each year have been delivered in government hospitals in recent years. It has been pointed out that 40 to 60 per cent of all women interviewed in maternity hospitals through the post-partum project actually availed themselves of the contraceptive services of the national program; and that these post-partum cases accounted for one-half of all its new acceptors. Furthermore, about 60 per cent of all acceptors attended the clinics within three months of the termination of their last pregnancies and about 80 per cent did so within one year. While such early practices after delivery undoubtedly imply some waste of contraceptive use because of post-partum infecundity, delay in acceptance of services could increase greatly the risk of unplanned and unwanted pregnancy.

In any case, the substantial acceleration in fertility decline which occurred simultaneously with the implementation of the national program points to program impact of a magnitude that must be far greater than any possible substitution functions would alone explain. In brief, there can be no doubt that the family planning program in Singapore had a major impact in accelerating the fertility decline which had previously begun.

The importance of the contribution made by the national program becomes even more apparent when considered in relation to the fertility prospect in Singapore. By reason of the fact that the postwar baby boom cohorts have now reached reproductive age, Singapore has experienced in the last two years a slight reversal in the downward trend in the birth rate. The birth rate increased from 21.8 in 1969 to 22.1 in 1970, and to 22.3 in 1971, in spite of the success of the national program. A study of the monthly registration of births shows that the birth rate probably rose again in 1972. Indeed, a projection of the population into the eighties (Saw Swee-Hock, 1970a) reveals that throughout the coming decade the proportion of women of ages 20-29 will continue to increase substantially. Other things being equal, this change in age composition will significantly push up the birth rate.

It has been demonstrated that nuptiality changes played an important role in the decline in fertility. In fact it was the major source of the decline before the inception of the national family planning program in 1966. However, precisely because the trend toward delayed marriage has already contributed so much to the decrease in fertility, much more reduction cannot be expected from this source. An international comparison of marital patterns reveals that, as a consequence of the recent trend of marriage postponement, age at first marriage in Singapore is currently among the highest in the world. It is higher than in many western countries where the "European marriage pattern" had previously prevailed, and in Asia only Japan and Hong Kong are comparable in this respect. It is therefore unlikely that age at marriage will rise much further in the coming years. The conclusion is inevitable, therefore, that further reduction in the birth rate, indeed, even the prevention of its continuing upsurge in the coming decade, hinges very much on the further reduction of fertility within marriage. Unless marital fertility continues to decrease, the birth rate of the population is likely to rise or at least remain relatively stable at a level above 20 per thousand throughout the seventies.

The government, however, is fully determined to bring the birth rate of the population to below 20 per thousand. The national family planning program has been extended for five more years, from 1971 to 1975. The target for this second five-year plan is to recruit 80,000 new acceptors with the purpose of reaching 60 per cent of the estimated 83,000 newly married women in this period and an additional 30,000 women from the "hard core" group, i.e., those women who were eligible for but did not adopt contraception during the first five years (Wan Fook Kee, 1972). The practice of child-spacing has been

very popular among the program's acceptors and many of them participated in the program before having their second or even first child. There is, moreover, evidence suggesting that the early practice of contraception for spacing purposes has, over the years, diffused from the upper to the lower strata of the population and thus has become more widely adopted. The evidence quoted by Chen (1972) also shows that one quarter of the so-called "hard core" women have not accepted contraception simply because they did not know of any method and 13 per cent because of lack of confidence in the effectiveness or safety of the methods they knew. The difficulties involved in such cases are presumably relatively easy to surmount. Thus, although in the nature of things it is not certain that the second five-year plan will achieve the target set for it, it can reasonably be expected that the national program will be able to reach a fair proportion of the "hard core" group and many women who will be newly married or at a very early stage in their family building ([11]).

Although Singapore has achieved considerable success in the dissemination of the practice of contraception among her people, she now faces perhaps an even more difficult challenge, that of reducing the desired family size of the population. In his study of the post-partum cases, D. Wolfers (1970b) found that for all three ethnic groups the number of children desired is between three and four. On the basis of response to a question on reason for family planning Chang and Yeh (1972) estimated of the program's acceptors in general that they want on the average 3.6 children. That the desire for three or four children is typical of the whole population is confirmed by the recent ECAFE study (Chen, 1972) which shows that 3.6 is the average desired number of children of a national sample of married women. Such a desired family size is admittedly only moderate by Asian standards; but it implies a rate of natural increase, given the low mortality and the already high population density of over 9,000 persons per square mile, that Singapore obviously cannot sustain for any long period.

If it is agreed that a family with three or four children is too large for Singapore, it is clear that a family planning program geared merely to increasing the demand for contraceptive services is by itself

([11]) The number of new acceptors of the national program in 1971 was 17,749, exceeding the annual target of 16,000 (Wan Fook Kee, 1972). It is commonly known, however, that many women, especially those of higher socio-economic status, prefer seeing private doctors for family planning services. The often long waiting time for service at the family planning clinics could be one reason. It is not known precisely how much this tendency may obstruct the substitution functions of the national program. The Board is aware of this and has tried to improve the clinic services.

inadequate to the solution of Singapore's population problem. With age at first marriage for women at an average of 23 years, the fertile period of more than twenty years after marriage is obviously long enough for the spacing of three or four children. In other words, even if Singapore is able to attain the ideal of universal and effective practice of family planning, a serious problem will still remain unless the desired family size is reduced. Thus, to tackle the problem of excessive population growth requires the creation of a two-child norm in addition to the further dissemination of birth control information and services.

The need to reduce family size ideals has not, of course, gone unheeded. Even before the national program was started, when family planning activities were still under the direction of the Association, the picture on posters and other publicity material portraying a couple with three children had been deliberately replaced by one of a family with only two. For the same reason, the central theme of the publicity and health education campaigns of the national program was in 1968 reoriented from "Plan your family" to the more definitive "Plan your family small". And, as mentioned above, the government has recently further formulated and revised social policies with the explicit purpose of discouraging parents from having more than two children. The evidence indicates that the desired family size of program acceptors may be declining. That this is the case is further suggested by the recent finding (Chen, 1972) that newlyweds prefer a smaller family on the average than does the married population as a whole.([12]) It remains to be seen how, in response to the implementation of the newly announced social policies, this trend may be accelerated and accentuated.

Although it is clear that the Singapore National Family Planning Program has significantly accelerated the reduction of fertility over the period studied, the longer run implications of its impact remain to be determined. The available data are all period, not, cohort data, and in consequence, it has not been possible to draw definitive conclusions on a cohort basis. This may constitute a serious limitation, especially for the younger cohorts including the postwar baby boom cohorts. The evidence that most acceptors are practicing family planning for child-spacing rather than child-limiting purposes leaves open the possibility that the number of children achieved by the younger cohorts will still be three to four rather than two as targeted by the government. Even should this occur, however, Singapore's growth rate would be reduced by reason of the increased length of

([12]) The desired family size of the newlyweds has been found to be 2.78 in comparison with 3.6 for the married population as a whole.

generation brought about by child-spacing; but such a decrease in the growth rate would not alone meet the government's objective.

In conclusion, the Singapore National Family Planning Program has certainly had a major impact in reducing fertility over the period studied and although the indications are that this impact will have significant longer-run consequences in the reduction of fertility and population growth, it is still too early to reach definitive conclusions about the future.

REFERENCES

Chang, C. T.

1974 Fertility Transition in Singapore. Singapore University Press.

Chang, C. T. and Stephen Yeh

1972 "A study of Singapore's national family planning programme", Malayan Economic Review, XVII, 1 (April).

Chen, Peter S. J.

1972 A Comparative Study of Husband-Wife Communication and Family Planning in Four Asian Countries: Singapore, Report No. 2, ECAFE, Bangkok.

Chow, L. P. and S. C. Hsu

1969 "A Chinese view of family planning in the developing world" in S. J. Behrman, L. Corsa Jr. and R. Freedman (eds.) Fertility and Family Planning: A World View, Ann Arbor, University of Michigan Press.

Chua, S. C.

1964 State of Singapore, Report on the Census of Population, 1957.

Del Tufo, M. V.

1949 Malaya. A Report on the 1947 Census of Population.

Freedman, R., A. Hermalin and T. H. Sun,

1972 "Fertility trends in Taiwan", Population Index, 38 (April-June).

Kanagaratnam, K. and C. K. Khoo

1969 "Singapore: The use of oral contraceptives in the national program", Studies in Family Planning, No. 48 (December).

Kanagaratnam, K. and Thonk Kah Leong

1970 "Role of post-partum program in the national program of Singapore" in Gerald I. Zatuchni, ed. Post-partum Family Planning: A Report on the International Program, ed. McGraw-Hill.

Kitagawa, Evelyn M.

1955 "Components of a difference between two rates", Journal of the American Statistical Association, 50, pp. 1168-1195.

Kwa S. B., S. T. Quah and Mark C. E. Cheng

1971 "The Abortion Act, 1969: A review of the first year's experience", Singapore Medical Journal, 12: 5 (October).

Lim Chong Yah and Ow Chwee Huay

1971 "The economic development of Singapore in the sixties and beyond" in You Poh Seng and Lim Chong Yah (eds.), The Singapore Economy, Singapore, Eastern Universities Press.

Ryder, N. B.

1957 "The conceptualization of the transition in fertility", Cold Harbor Symposia on Quantitative Biology, 22.

Saw, Swee-Hock

1970a Singapore Population in Transition, University of Pennsylvania Press, Philadelphia.

1970b "Estimation of Singapore's population by quinary age-groups for 1962 and 1967", Malayan Economic Review, XV (April).

Singapore

1971 Report on the Registration of Births, Deaths and Marriages, 1966-1970.

Singapore Family Planning and Population Board

Annual Report, 1966-1970, Singapore.

United Nations

1965 Population Bulletin No. 7—With Special Reference to Conditions and Trends of Fertility in the World, New York.

Wan Fook Kee

1972 "Singapore", Studies in Family Planning, Vol. 3, No. 7 (July).

Wan Fook Kee and S. T. Quah

1971 "Singapore: A study of clinic continuation rates", Studies in Family Planning, Vol. 2, No. 12 (December).

Wolfers, D.

1969 "The demographic effects of a contraceptive program", Population Studies, 23 (March).

1970a "The Singapore family planning program: further evaluation data", American Journal of Public Health, Vol. 60, No. 12 (December).

1970b "The epidemiology of acceptors of post-partum intra-uterine contraception" in D. Wolfers, S. H. Tow and S. S. Ratnam (eds.), Post-Partum Intra-Uterine Contraception in Singapore, Excerpta Medica Monographs Series, Amsterdam.

Wolfers, H.

1970 "The social and psychological correlates of acceptance" in D. Wolfers, S. H. Tow and S. S. Ratnam (eds.), Post-Partum Intra-Uterine Contraception in Singapore, Excerpta Medica Monographs Series, Amsterdam.

You, Poh Seng

1967 "The population of Singapore, 1966: demographic structure, social and economic characteristics", Malayan Economic Review, XII No. 2 (October).

You, Poh Seng *et al.*

1967 Singapore Sample Household Survey, 1966, Report No. 7.

You, Poh Seng, V. V. Bhanoji Rao and G. Shantakumar

1971 "Population growth and population characteristics" in You Poh Seng and Lim Chong Yah (eds.) The Singapore Economy, Eastern Universities Press, Singapore.

Chapter 12

THE IMPACT ON FERTILITY OF TAIWAN'S FAMILY PLANNING PROGRAM

by
T. H. Sun

I. INTRODUCTION

The purpose of this chapter is to describe and discuss the systematic analyses undertaken in Taiwan in studying the impact of its family planning program on fertility. Considerable skepticism about the impact of family planning programs on a nation's fertility has been expressed by some demographers (Davis, 1967; Hauser, 1967). Taiwan has been a focus of international attention in this respect since its highly accurate vital statistics and relatively complete family planning service statistics make it the most likely candidate for the establishment of a cause and effect relationship between program effect and declining fertility. A series of closely evaluated studies have been carried out during the past decade in order to document the nature of this relationship. The picture, however, has been complicated by the fact that the program was started during a period of declining fertility.

The study area of the chapter includes Taipei City and Taiwan Province of the Republic of China. This island, which is about 245 miles long and 85 miles wide at its widest point, has one of the highest population densities in the world—about 1,100 persons per square mile (in 1973) in an area of 14,000 square miles. The Chinese population of Taiwan increased from a few score of hundreds in 1600 to an estimated 50,000 by the 1660s; in the 1680s there were estimated to be 120,000 Chinese on the island and this number had increased to about 2 million by 1811. The island was ceded to Japan in 1895. The first census conducted in 1905 gave a population estimate of 3,123,302 and by 1943 this number had increased to 6,585,841. When the Japanese surrendered in 1945, Taiwan was restored to the Republic of China. In 1946 about 500,000 Japanese returned to Japan but during 1947-1951 more than a million Chinese came to Tai-

Indicators	1952	1956	1960	1964	1968	1972
Economic indicators						
Gross national product index (a)	100	133	171	252	360	535
Per capita real national income index (a)	100	111	125	166	203	272
Savings as percentage of national income	4.8	4.9	6.9	15.6	20.9	27.6 (b)
Agricultural production index	100	121	143	178	226	247
Total industrial production index (c)	100	155	245	410	803	1,701
Electric power production index	100	158	255	419	690	1,229
Per cent of net domestic product from industry	18	22	25	28	30	37
Per cent of male labor force over age 11 in nonagricultural occupation	42	45	49	50	60 (d)	67 (d)
Education indicators						
Per cent of population aged six and over who are illiterate	42	37	27	22	16	13
Per cent of primary school graduates enrolled in junior high school	34	47	51	55	75	84
Communication and transportation indicators						
Daily newspaper circulation per 100 population	28	40	55	70	76 (f)	83 (f)
Radios per 1,000 families	34	105	350	555	629 (e)	650 (f)
Posting of domestic letters per capita	8	17	33	32	32	45
Health indicators						
Death rate per 1,000 population	10	8	7	6	5	5
Life expectancy at birth for males	56	60	62	64	64	66
Population per doctor	1,610	1,537	1,553	1,520	1,438	1,306

(a) At constant price of 1966 and adjusted for gain or loss due to changed terms of trade.
(b) 1971 data.
(c) Includes manufacturing, mining, electricity, gas and water.
(d) For male labor force over 14 years of age.
(e) 1969 data.

wan from the mainland. The population size doubled in only 18 years, reaching 12,256,682 in 1964. At the end of 1973 it was 15,564,830.

The population of Taiwan is relatively young due to the high birth rates of the past. The proportion of the population aged under 15 increased from 43.4 per cent in 1955 to 46.0 per cent in 1962 but had decreased to 37.9 per cent by 1972. On the other hand, the proportion of women in the peak fertility period (20-29) amongst all women of child-bearing age (15-49) had increased from 30.7 per cent in 1969 to 32.9 per cent in 1972. It is projected that this proportion will increase to 40.3 per cent by 1980. Fortunately, women are marrying later than previously, the mean age at marriage having increased from 19.0 years in 1940 to 20.9 in 1948, 21.9 in 1963 and 22.6 in 1967. The proportions married at ages 15-19 and 20-24 have declined from 11.4 and 59.3 per cent in 1963 to 6.6 and 46.4 per cent respectively in 1972.

The literacy rate in Taiwan is one of the highest in Asia; in 1972, 91 per cent of the males and 72 per cent of the females aged over 15 were literate. Since 1968 compulsory education has been increased to nine years (including junior high). In the 1972-1973 school year, 84 per cent of the primary school graduates enrolled in junior high schools, 71 per cent of the junior high school graduates in senior high schools and 83 per cent of senior high school graduates in schools of higher education.

There is little external migration but considerable movement towards the cities and industrialized areas. Taipei and Kaohsiung, the two largest cities, attract most of the internal migrants. The average annual rates of increase by migration in these cities between 1970 and 1973 were 21 and 23 per thousand, respectively.

The proportion of the male labor force over age 15 engaged in non-agricultural occupations has risen from 42 per cent in 1952 to 64 per cent in 1972. Some indicators of economic and social development are shown in Table 12.1. They show that rapid changes in a wide range of areas are taking place in Taiwan. As James Grant (1972) has pointed out, not only has Taiwan's increase in national product been spread broadly over both the rural and urban populations, but income distribution has actually improved markedly. The poorest 20 per cent of both the rural and urban populations today receive a significantly larger share of the greatly increased national income than they did 20 years ago. It is well known that significant fertility declines are likely to be associated with such major changes in the social and economic structure.

THE BEGINNINGS OF THE DEMOGRAPHIC TRANSITION

It is clear from Figure 12.1 that Taiwan is in the second stage of its demographic transition, i.e., mortality has been low for some time and fertility has also been falling (Freedman, Takeshita and Sun, 1964).

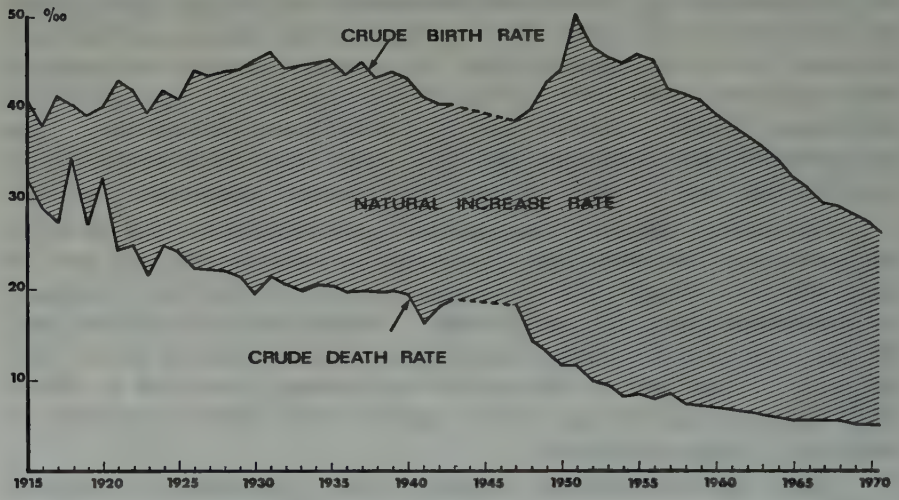


Fig. 12.1. *Vital rates in Taiwan area, 1915-1972*

The crude death rate in Taiwan fluctuated around 30 per thousand until 1920, after which it decreased steadily to below 20 per thousand in 1945. After the Second World War it declined dramatically from 18 per thousand in 1947 to 8 per thousand in 1956 as a result of the improvement in public health, the increased use of modern drugs and the general improvement in living standards; by 1972 it had reached 4.8 per thousand. As might be expected, infant mortality is also low, about 27 per thousand (corrected for underregistration). These statistics provide reasonably accurate assurance that most children born will survive to adulthood.

On the other hand, the crude birth rate remained quite high, around 40 per thousand, until the end of the Second World War and after the War it even rose from 38 per thousand in 1947 to 50 per thousand in 1951. However, a steady decline set in in 1958, six years before the island-wide expansion of the family planning program and by 1972 it had reached 24.2 per thousand.

THE EXTENSIVE FAMILY PLANNING PROGRAM

The family planning activities in Taiwan were started by the China Family Planning Association in 1954 in the form of training in "first-aid" for women in military dependents' villages and some rural areas. This activity did not develop into a large scale program, however, due to the size and non-official status of the organization, the limitation of its budget and staff and the lack of a national policy.

In 1959 the Taiwan Provincial Health Department started to provide pre-pregnancy health (PPH) services in one of the Island's 22 counties (with a population of 400,000) as part of the government's Maternal and Child Health (MCH) program. The Governor of Taiwan Province was convinced that a reduced rate of population growth would act as a stimulus to the economy. This program was later expanded (1963) to cover 120 of the 361 local health stations. One full-time PPH worker ([1]) was employed in each of the 120 health stations to motivate women to use conventional contraceptive methods.

In 1964, after a careful experimental program action in Taichung City (Berelson and Freedman, 1964; Freedman and Takeshita, 1969), this PPH program was strengthened by the introduction of the Lippes loop, an intra-uterine device, as the main contraceptive method. A five-year target of 600,000 loop insertions was set in an attempt to bring the population growth rate down from 3 per cent to 2.0 per cent over the five years (1964-1969), and to below 2 per cent in ten years (by 1973). In 1964, about 200 private doctors were trained to insert loops and between 1964 and 1967 the program spread to 331 of the 361 areas. The number of PPH workers increased gradually to more than 300 and the number of doctors inserting loops to 680. Oral contraceptive pills were introduced in 1967 to supplement the loop. More than eight per cent of the currently married women in the child-bearing ages (20-44) ([2]) accepted one of these two methods in a single year (1967); there were 121,053 loop insertions and 24,878 pill acceptors. Behind this expansion of the program was the strong backing of the Minister of Economic Affairs who stressed the economic advantages of reducing the population growth rate from 3 to 2 per cent and made US $1.5 million available for the operation of the program during the 1964-1969 period.

With the promulgation of the "Regulations Governing the Implementation of Family Planning in Taiwan Area" in May 1968, the family

([1]) Full-time home-visiting field worker whose main job was to introduce contraceptive methods to couples who needed them.

([2]) Ages 20-44 were used in Taiwan due to the rarity of early marriages. (The proportion married for ages 15-19 is only 8 per cent).

planning program was given an official status. Since then, the education component has been strengthened and there has been more participation of organizations other than those concerned with health. The first five-year plan was completed in 1969, and the second begun in 1972. The total budget for the program increased from US $300,000 in 1965 (2.3 cents per person) to about US $1 million in 1972 (about 7 cents per person). Starting in 1970, condoms were provided free by the program. The number of full-time field workers increased to 450 in 1972 and the number of loop-insertion doctors to 895. Family planning is now regarded as one of the most important programs in the Provincial Government administration and is listed as one of the two important special programs of the Health Department put under strict supervision and evaluation by the central government.

THE AVAILABILITY OF HOUSEHOLD REGISTRATION DATA

The Taiwan Area has an excellent household registration system which provides data for evaluating the effect of family planning on fertility.

The system was set up in 1905 by the Japanese for better administrative control and was revised by the Chinese government after the Second World War. Each township has a household registration office, which keeps the household records of people in the township and periodically makes basic tabulations.

Each household has a household booklet which records: address, names of the household members, their relationship to the head of the household, order of birth, personal identity (ID) number and date of birth. The local police are responsible for checking the households monthly to press for the reporting of any changes which have occurred, such as childbirth, death or migration. The reporting of these events remains the responsibility of the head of the household.

All persons aged 18 years or more should have an ID card which records their ID number, date of issue, name, date, order and place of birth, names of parents and spouse, level of education, occupation, and address.

The static data registered at the household registration office include: address (both administrative division and street number), name of head of the household, and for each person who lives in the household: name, ID number, his (or her) relationship to the head of the household, date and order of birth, level of education, occupation, names of parents and spouse, and place of origin. Any change in the items listed above will be recorded when it is reported. These data are tabulated annually (at the end of each year) by town-

ship. The tabulations include the number and size of households, size of population by current address and by place of original domicile, sex, age composition, level of education, marital status, and, for those aged 15 and above, employment status, occupation and industry. Of these data, age and sex are fairly accurate; but for other items accuracy is lower because of the delay in reporting changes. The accuracy rates range from about 70 to 90 per cent depending on the item.

The dynamic data registered are: births, deaths, changes of address, marriages, and adoptions. The numbers of each of these items are tabulated and reported monthly by township. The data on births are further tabulated quarterly by age of parents, and those on deaths by age at death. Birth registration is fairly accurate apart from some delay in reporting. Death registration is also quite accurate except for infant mortality, which is estimated to be underreported by about 45 per cent (Sullivan, 1972).

In summary, the household registration system provides the following data which can be used for program planning and evaluation:

(1) Monthly:

a) Total population.
b) Number of births.
c) Number of deaths.
d) Names and addresses of women who have had a live birth during the month.

(2) Annually:

a) Population by age, sex, marital status, education, and occupation.
b) Births by age of mother.

THE PROGRAM RECORD SYSTEM AND ORGANIZATIONAL SET-UP FOR EVALUATION

Taiwan's family planning program was very much research-oriented from the beginning. Even before the expansion of the loop program to the whole island in 1964, a carefully designed experimental action program was carried out by the Taiwan Population Studies Center and the Mother and Child Health Institute in Taichung City in 1963 with the assistance of the University of Michigan and the Population Council. The operation of this experimental program was carefully recorded and analyzed together with the "before" and "after"

surveys to provide the basic data necessary for expansion of the program (Berelson and Freedman, 1964; Freedman and Takeshita, 1969).

When the program was expanded in 1964, the Taiwan Population Studies Center continued to act as its evaluation unit; the unit analyzed the program input records and conducted evaluation surveys and action-oriented experimental programs to find out how well the program was meeting its objectives, what were the specific reasons for its success or failure, and how its effectiveness and efficiency could be increased. The results of these studies were used to guide the program. This organizational evaluation set-up was necessitated by the fact that none of the workers in the program had had experience of a large-scale family planning program and that a successful operation was essential to convince the economic planners and finance people of the importance and benefits of the program, which was quite costly despite its unofficial status.

The service statistics collected from the program include:

(1) Program inputs: budget and actual expenditure, manpower, number of clinics, volume of contraceptive supplies and educational materials distributed, number of home visits made, number of meetings organized and number of people who attended them.
(2) Program outputs: number and characteristics of acceptors, results of IUD and pill follow-up surveys, changes in the level of knowledge, attitude and practice of contraception (KAP) and changes in fertility levels.

The program input data are obtained mainly through a record-keeping system at headquarters combined with regular reporting from the field. The coupon system has been the main source of information regarding program method acceptances. This system provides prompt information on the number and characteristics of acceptors of a specific method by local area and referral agent (Cernada and Chow, 1969). The returned coupons have also been used as the sampling base for follow-up surveys on particular contraceptive methods. So far, three island-wide IUD acceptor follow-up surveys (1965, 1967, 1969) and a series of follow-up surveys of the 10,000 IUD acceptors in the Taichung family planning clinic have been carried out. For pill acceptors, an island-wide survey was conducted in 1968, followed by a study to find out why the continuation rate was so low. A follow-up survey of condom acceptors is currently underway. Four island-wide KAP surveys, which also contain detailed information on fertility history, have been carried out (1965, 1967, 1970 and 1973).

A survey of economic correlates of fertility was carried out in 1969. The program has been evaluated at three levels:

(1) Monthly evaluation of program achievements compared with program targets: service statistics and returned coupons.
(2) Evaluation of specific contraceptive methods: follow-up surveys of acceptors.
(3) Evaluation of the program effect: KAP surveys, household registration data and cost-benefit analyses.

Most of the mid-term and long-term evaluation activities are currently carried out by the Survey and Research Division of the Committee on Family Planning (of the Provincial Department of Health) which has an efficient data processing unit equipped with an IBM 1130 computer. The monthly evaluation is carried out by the Planning Division of the Committee. The research and evaluation activities have been facilitated by the assistance of the Population Studies Center of the University of Michigan in the planning of surveys and the analysis of survey data. The Population Council has also provided financial and technical assistance, not only for the program evaluation but also for its operation. The National Health Administration of the Central Government began to carry out some overall evaluation research activities in 1972.

II. PROGRAM ACCEPTANCES: QUANTITY AND QUALITY

The primary objective of the family planning program in Taiwan was to enable married couples to practice contraception for spacing their childbirths and preventing unwanted pregnancies. The objectives have recently been broadened to include more motivational activities aimed at reducing the ideal family size and lessening the degree of preference for sons, which are the two important obstacles to further fertility decline. The effectiveness of the program can therefore be measured by indicators such as the number or rate of acceptances, the characteristics of the acceptors and the use-effectiveness of the contraceptive methods accepted.

ACCEPTANCE RATE

When the first island-wide KAP survey was carried out in 1965, only 22.8 per cent of the married women aged 20-44 were currently

using contraception and only 26.6 per cent had had any experience of contraceptive use. One of the objectives of the program was, of course, to increase the practice rate by motivating eligible couples to accept effective contraceptive methods. When the loop was introduced into the island-wide program in 1964, the coupon system was also established to collect acceptance data (Cernada and Chow, 1969). One coupon per acceptor was submitted to headquarters in the month following the acceptance. These coupons were numbered and tabulated by townships and fieldworkers, and were compared monthly with the targets set for these local areas. At present, all information on coupons is punched on IBM cards and tabulated monthly by computer. When pills and condoms were introduced into the program, the same system was employed to collect the information needed for evaluation of these two methods. A 5 per cent sample is randomly selected from the coupons at about three-monthly intervals to check the reliability of the field supervisor's reports on acceptance. So far, the reporting error has been less than 5 per cent. It is possible, however, that the 15 per cent of the sample who could not be contacted might have been subject to a higher rate of reporting error.

As shown in Table 12.2, by the end of 1971, 762,912 married women aged 15-54 had had IUD (Lippes loop) insertions and including reinsertions, a total of 934,715 insertions had been made. In addition, 229,068 women had accepted pills and 110,415 couples had accepted condoms. The number of first loop insertions has been maintained at around 100,000 women per annum since 1965. The annual rate of first insertions per 100 married women aged 20-44 years has also been stable at around 6 with a slightly declining trend, but if the rate were calculated on the base of women eligible for first insertions, then it has increased from 6.3 per cent in 1965 to 8.7 per cent in 1971. These facts indicate that program acceptance in Taiwan has not yet reached the plateau, if one exists, which is supposed to indicate that all of the "elite" ready group has been reached. By the end of 1971, a total of 39.6 per cent of the 1.8 million married women aged 20-44 had tried the loop at least once ([3]) (Table 12.3). This is a significant achievement having an important impact on the fertility of acceptors. The program, however, reaches a disproportionate number of older women. Table 12.3 shows that the cumulative proportion of women who had had a first IUD insertion in 1971 is as high as 50 per cent for those in the age-group 30-39. This is consistent with the faster decline of fertility among women in this age-group. Particularly considering that acceptors have had high past

([3]) Of whom approximately 90 per cent had never used any method before.

Year	Annual				Cumulative	
	No. of first IUD insertions (a)	Number of reinsertions (a)	Total insertions (a)	Percentage of reinsertions	Number of first IUD insertions (a)	Total insertions (a)
1964	49,496	754	50,250	1.5	49,496	50,250
1965	96,474	2,779	99,253	2.8	145,970	149,503
1966	105,680	5,562	111,242	5.0	251,650	260,745
1967	110,530	10,523	121,053	8.7	362,180	381,798
1968	101,533	22,137	123,670	17.9	463,713	505,468
1969	95,031	35,327	130,358	27.1	558,744	635,826
1970	99,024	44,281	143,305	30.9	657,768	779,131
1971	105,144	50,440	155,584	32.4	762,912	934,715

Year	Estimated number of married women 20-44 (midyear) (b)	Estimated number of married women 20-44 eligible for a first insertion (c)	Annual rate of first insertions per 100 married women 20-44 (d)	Annual rate of first insertions per 100 married women 20-44 eligible for a first insertion	Cumulative rate of first insertions per 100 married women 20-44 in this or any previous year (e)	Number of pill acceptors	Number of condom acceptors
1964	1,532,512	1,532,512	—	3.2	3.2	—	—
1965	1,574,240	1,525,729	6.1	6.3	9.0	—	—
1966	1,607,968	1,465,393	6.5	7.1	14.7	—	—
1967	1,634,776	1,390,559	6.7	7.9	20.3	27,552	—
1968	1,667,418	1,318,806	6.0	7.6	24.9	35,622	—
1969	1,707,653	1,265,606	5.5	7.4	28.6	32,208	—
1970	1,748,474	1,222,443	5.6	8.0	32.1	54,517	49,121
1971	1,790,144	1,184,871	5.7	8.7	35.5	79,169	61,294

(a) The figures in these columns include insertions (about 1 per cent of the total) to women 45 years of age or older and to women 19 years of age or younger.
(b) From Taiwan Demographic Fact Books.
(c) Excludes women who had IUD's inserted in a previous year and are in the age group 20-44 in this year, as well as women who passed out of the age range. No allowance has been made for the small number of cases affected by divorce or death.
(d) The rates in this and the next two columns exclude the estimated number of acceptors under 20 and over 45 years of age.
(e) The cumulative percentage of women aged 20-44 in this or any previous year.

Source: Freedman *et al.*, 1972, p. 154.

Table 12.3. *Cumulative number and percentage of married women who had a first IUD insertion, 1964-1971 (a), Taiwan area*

Age	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*	*1971*
				Cumulative number				
15-19	99	384	709	1,072	1,438	1,858	2,422	3,183
20-24	3,131	9,571	16,755	23,429	29,345	35,565	43,107	53,438
25-29	12,733	36,686	61,347	84,621	102,647	116,539	129,917	144,861
30-34	17,031	48,242	79,972	112,316	139,747	163,276	186,659	208,728
35-39	12,140	35,411	60,760	87,706	113,486	138,208	161,457	184,794
40-44	4,053	13,850	26,889	42,717	59,954	78,075	97,874	116,228
45-49	309	1,764	4,953	9,663	15,860	23,231	32,684	44,684
50-54	—	62	265	656	1,236	1,992	3,648	6,996
20-44	49,088	143,760	245,723	350,789	445,179	531,663	619,014	708,049
15-54	49,496	145,970	251,650	362,180	463,713	558,744	657,768	762,912
				Cumulative rate of acceptance (per cent)				
15-19	0.2	0.8	1.3	1.8	2.2	2.8	3.6	4.9
20-24	1.1	3.6	6.4	9.2	11.6	13.5	15.5	17.6
25-29	3.3	9.4	15.2	20.8	25.0	28.3	32.2	37.1
30-34	4.9	13.4	21.7	29.8	36.1	40.5	45.8	50.0
35-39	4.7	11.4	18.9	26.4	33.4	39.1	45.2	50.6
40-44	1.8	5.7	10.5	16.6	21.7	26.4	32.4	37.0
45-49	0.2	1.0	2.6	4.8	7.5	10.2	14.0	18.3
50-54	—	0.0	0.2	0.4	0.8	1.2	2.2	4.0
20-44	3.2	9.1	15.3	21.5	26.7	30.8	35.4	39.6
15-54	2.6	7.5	12.6	17.6	22.1	25.9	29.7	33.6

(*a*) Insertions were made at specific age or a younger age in specific year or earlier year. The cumulative age distribution of acceptors was estimated by advancing one-fifth of the acceptors in each age group to the next older age group in each year. This estimation procedure was necessary since the data were not available for single years of age.

fertility, their contraceptive use in the program should have reduced fertility in the ages involved.

Of the 225,048 women who accepted pills through the program from 1967 to 1971, 63,700 (about 3.5 per cent of the 1.8 million married women aged 20-44 in 1971) had had no previous experience of any contraceptive method.

Of the 110,814 new condom acceptors in the program from 1970 to 1971, 36,000 (a further 2 per cent of the 1.8 million married women aged 20-44 in 1971) had had no previous contraceptive experience.

Taking loop, pill and condom together, roughly 45 per cent of the eligible couples had been recruited by the program by the end of 1971. This fact indicates that the program has provided a major means for the adoption of birth control leading to fertility decline. Substitute means might have been used in its absence, but that is not certain. What is certain is that a large number of couples reduced their fertility after contact with the program.

CHARACTERISTICS OF ACCEPTORS

The impact of program acceptance on fertility depends partly on the type of acceptors recruited. The acceptor characteristics having an impact on fertility for which information is available for Taiwan are: age, parity, contraceptive experience before acceptance, length of open birth interval, purpose of acceptance, and educational background.

As shown in Table 12.4 the mean age of loop acceptors in the Taiwan area is relatively high: a median age of 32.6 in 1964 and 30.1 in 1971. The proportion of loop acceptors younger than age 30 was only 32 per cent in 1964. This proportion, however, had increased to 49 per cent by 1971. About 70 per cent of the loop acceptors in 1964 had already had four or more children and most of them (84 per cent in 1966) wanted no more children. This was to be expected as it was felt in Taiwan that the family planning program, given its limited resources, should begin with the "ready" group, which was easier to approach and which would have an impact in preventing excess fertility. The emphasis was later changed, however, to focus on spacing for the younger age groups with fewer children. This is believed to have increased the program acceptance and hence the impact on fertility. The proportion of loop acceptors with three or fewer children increased from 31.7 per cent in 1964 to 54.7 per cent in 1971, and the proportion of women accepting the loop for the spacing of children increased from 15.8 per cent of all acceptors in 1966 to 20.6 per cent in 1971. This change is important

Table 12.4. *Percentage distribution of IUD acceptors in family planning program by age, number of living children, education contraceptive experience, and purpose of using IUD, 1964-1971, Taiwan area*

Characteristics	1964	1965	1966	1967	1968	1969	1970	1971	Total	
									Number	Percentage
Age										
24	6.5	7.5	8.7	9.2	10.5	12.6	15.0	17.7	108,826	11.8
25-29	25.6	26.6	28.3	29.3	30.3	31.5	31.8	31.6	275,809	29.9
30-34	34.3	33.2	32.0	31.8	30.7	29.3	28.1	26.6	277,795	30.2
35-39	24.6	23.3	21.5	20.3	19.0	17.5	16.5	15.5	174,925	19.0
40-44	8.3	8.7	8.5	8.4	8.3	8.0	7.4	7.3	73,936	8.0
45 and above	0.7	0.8	1.0	1.0	1.2	1.1	1.2	1.3	9,998	1.1
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	921,289 [a]	100.0
Median age	32.6	32.4	32.0	31.8	31.5	31.0	30.6	30.1		31.4
Number of living children										
0	0.1	0.3	0.2	0.1	0.1	0.2	0.2	0.2	1,525	0.2
1	1.7	2.3	3.1	3.1	3.9	4.7	5.2	6.6	37,892	4.2
2	9.2	10.7	12.6	13.4	15.2	16.4	17.3	19.5	136,611	15.1
3	20.7	21.4	22.9	24.3	25.5	27.2	28.1	28.4	230,488	25.6
4	25.6	25.0	24.8	25.0	24.5	23.9	23.8	22.7	218,278	24.2
5	19.4	18.8	17.2	16.5	15.2	14.2	13.4	12.2	137,378	15.2
6 or more	23.3	21.5	19.2	17.7	15.6	13.4	12.0	10.6	140,041	15.5

				40.4	37.1	34.8	31.6	28.6	288,169	36.0
Primary	44.3	45.3	48.0	49.0	51.1	52.4	54.9	56.7	410,802	51.3
Jr. high	8.7	7.7	7.1	6.7	7.2	7.4	7.5	8.0	59,518	7.4
Sr. high or more	5.7	4.7	4.5	4.3	4.6	5.4	6.0	6.8	42,342	5.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	800,831 (c)	100.0
Contraceptive experience										
Never used before			83.0	78.2	71.6	64.3	61.5	59.2		
Used loop			5.2	9.8	15.3	19.7	23.2	23.4		
Used pill			0.7	1.0	1.8	3.0	3.4	3.9		
Used other methods			11.1	11.0	11.3	13.0	11.9	13.5		
Total			100.0	100.0	100.0	100.0	100.0	100.0		
Purpose of acceptance										
Spacing			15.8	15.1	16.9	18.3	19.5	20.6		
Stopping			84.2	84.9	83.1	81.7	80.5	79.4		
Total			100.0	100.0	100.0	100.0	100.0	100.0		

(*a*) Excluding 9,776 cases whose age is unknown.
(*b*) Excluding 28,852 cases whose parity is unknown.
(*c*) Excluding 130,234 cases whose education is unknown.
Source: Taiwan Provincial Committee on Family Planning, 1972.

Table 12.5. *Percentage distribution of pill acceptors in family planning program, by purpose of using pills, and contraceptive experience before acceptance, 1967-1971, Taiwan area*

Purpose and experience	*1967*	*1968*	*1969*	*1970*	*1971*
Purpose of acceptance:					
Spacing	8.1	10.5	14.2	16.4	20.0
Stopping	91.9	89.5	85.8	83.6	80.0
Total	100.0	100.0	100.0	100.0	100.0
Contraceptive experience:					
Never used before	10.1	16.8	19.5	32.0	40.0
Used loop	65.0	58.2	53.3	49.8	39.5
Used pill	5.7	6.2	8.1	8.7	9.4
Used other methods	19.2	18.8	19.1	9.5	11.1
Total	100.0	100.0	100.0	100.0	100.0

Source: Committee on Family Planning.

Table 12.6 *Contraceptive experience of pill acceptors in family planning program, by education and number of living children, 1970, Taiwan area (percentages)*

Education and parity	*Contraceptive experience before acceptance*					
	Never used	*Used loop*	*Used pill*	*Used others*	*Total*	
					%	*N*
Education:						
No formal	26.9	57.2	5.6	10.3	100.0	14,009
Primary	32.4	50.0	8.7	8.8	100.0	30,329
Jr. high	35.1	40.7	13.2	11.0	100.0	4,382
Sr. high or more	41.6	32.2	16.1	10.1	100.0	3,854
Total	31.8	49.8	8.8	9.6	100.0	52,574 (a)
No. of living children:						
0	79.4	5.8	12.0	2.8	100.0	773
1	66.6	20.7	8.9	3.8	100.0	3,436
2	42.4	40.7	10.0	6.9	100.0	9,288
3	28.9	52.2	9.1	9.8	100.0	14,112
4	24.7	55.9	8.2	11.2	100.0	12,254
5 or more	23.4	57.6	7.7	11.3	100.0	14,223
Total	31.8	49.8	8.8	9.6	100.0	54,086 (b)

(a) Excluding 2,481 cases for whom educational level is not available.
(b) Excluding 969 cases for whom number of living children is not available.
Source: Committee on Family Planning, 1971.

because it shows that the program is now recruiting more fecund young women who have a longer fertile period ahead. This is probably one of the reasons why Taiwan has been able to recruit about the same number of new acceptors every year for the past few years without "skimming the cream".

The characteristics of the pill and condom acceptors are quite similar to those of the loop acceptors, the proportions of spacers in 1971 being 20.0 per cent (Table 12.5) and 21.2 per cent respectively. The proportions of acceptors having three children or less were 56.0 per cent and 56.9 per cent for pill and condom acceptors respectively and their median ages 29.5 and 30.7.

One criticism of family planning programs is that the majority of their recruits are often those couples already using some form of contraception and very little is gained by transferring a couple from a non-program to a program method. This, however, has not been the case in Taiwan. According to the coupon returns, only 17 per cent of the loop acceptors in 1966 had had experience of other contraceptives, a proportion which increased to 41 per cent in 1971 due to the increase of loop reinsertions. Twenty-three per cent of 1971 loop acceptors had used the loop before (Table 12.4) but among those having a first insertion only 11 per cent had had previous contraceptive experience. This situation for pill and condom acceptors, however, started somewhat differently because these two methods were introduced later to supplement the loop. This was especially true of the pill which was offered only to loop dropouts in 1967-1970. Condoms were introduced in 1970 for all who wanted them, and since 1970, the pill has also been available to all eligible couples. Therefore the number of couples who accept the pill or condom as their first contraceptive method is increasing and about 32 per cent of the pill and 26 per cent of the condom acceptors in 1970 had had no previous contraceptive experience (Tables 12.6 and 12.7).

Tables 12.6 and 12.7 also show that the proportion never having used a contraceptive before accepting these two methods is much higher for those who are better educated or have fewer children.

Another important characteristic of acceptors is their educational level. All studies on the use of contraception show that the proportion of contraceptive use is higher among those with higher education. This is natural because better-educated people have more knowledge about contraceptives and are faster in adopting new things. Therefore, before the introduction of a large-scale family planning program, differentials of contraceptive use by education are usually very high, a fact which leads some people to think that the practice of family planning will produce "reverse selection". Since one of

Table 12.7. *Contraceptive experience of condom acceptors in family planning program, by education and number of living children, 1970, Taiwan area (percentages)*

Education and parity	Contraceptive experience before acceptance				Total	
	Never used	Used loop	Used pill	Used others	%	N
Education:						
No formal	20.4	48.3	10.4	20.9	100.0	8,048
Primary	25.8	41.3	10.5	22.4	100.0	26,371
Jr. high	28.0	33.0	11.0	28.0	100.0	7,266
Sr. high or more	33.9	27.9	11.0	27.2	100.0	7,116
Total	26.4	39.2	10.6	23.8	100.0	48,801 (a
No. of living children:						
0	71.3	3.2	5.9	19.6	100.0	697
1	62.9	16.1	6.7	14.3	100.0	3,490
2	37.3	31.9	10.6	20.1	100.0	8,647
3	22.7	41.5	12.1	23.7	100.0	13,469
4	18.3	45.1	10.7	25.9	100.0	11,273
5 or more	16.9	41.6	10.3	31.2	100.0	11,629
Total	26.4	39.3	10.6	23.7	100.0	49,203 (b

(*a*) Excluding 672 cases for whom educational level is not available.
(*b*) Excluding 270 cases for whom number of living children is not available.
Source: Committee on Family Planning, 1971.

Table 12.8. *Demographic characteristics of IUD acceptors by educational level, 1971, Taiwan area*

Characteristics	Level of education				
	No formal	Primary	Junior high	Senior high or more	Total
1. Median age	33.6	29.0	28.0	27.9	30.1
2. Median no. of living children at 1st insertion	4.0	3.1	2.3	1.9	3.1
3. Percentage having 3 or fewer children	32.2	58.1	78.6	89.7	53.7
4. Percentage having no sons	3.1	7.8	14.1	21.1	7.8
5. Mean open birth interval (in months)	30.6	20.4	17.9	14.2	22.7
6. Percentage using IUD for spacing	7.3	22.0	36.7	50.4	20.6

Source: Taiwan Provincial Committee on Family Planning, 1972.

the objectives of a family planning program is to introduce, through a massive effort, contraceptives to those who would not ordinarily adopt them, it should recruit more people from lower-status groups to prevent this "reverse selection" procedure. As shown in Table 12.4, 87 per cent of the loop acceptors had only primary school education or less, although only 79 per cent of women aged 15 years old and above fell into this educational category. According to the sample tabulations ([4]) of married women aged 20-44 by educational level from household registration data for 1966 and 1971, the loop acceptance rates (including reinsertions) for these two years and for 1964-1971 (cumulative) were as follows:

Education	*1966 %*	*1971 %*	*1964-1971 % cum.*
No formal education	6.8	6.7	50.8
Primary education	7.0	10.1	54.9
Junior high education	7.2	8.7	48.3
Senior high and above	5.8	7.9	37.5
Total	6.9	8.6	51.6

It is clear from the cumulative rates that the loop acceptance rate is higher among women with primary education and with no formal education than among those with higher education. In other words, the program has been effective in promoting contraceptive use among people with lower educational levels, many of whom would not otherwise have adopted contraceptives at an early date. Therefore, the program can be considered to have had an extra impact on fertility by recruiting low-education people.

The educational level of pill acceptors is the same as that of the loop acceptors, while that of the condom acceptors is higher; wives of 30 per cent of the condom acceptors in 1970 had junior high school education or better.

When loop acceptors are tabulated by educational level, it is clear that those with more education are also younger and have fewer children, a higher proportion have no sons and more of them are

([4]) The tabulations were made by the Provincial Department of Civil Affairs based on the 56 (for 1966) and 51 (for 1971) sample townships' household records. A minor adjustment was made to match the educational classification in these tabulations with that of the acceptance records.

accepting the loop within a short time after childbirth and for spacing purposes than those with lower education (Table 12.8).

Another factor related to modernization is urban-rural residence. The acceptors of IUDs in the cities are younger and of lower parity than in the rural areas (Table 12.9). However, unpublished data indicate that rural areas "catch up" with urban areas after a time-lag of about 5-7 years.

The length of the interval between the last childbirth and the date of acceptance also influences the impact on fertility. Ideally contraceptive use should start before resumption of ovulation, which occurs

Table 12.9. *Percentage distribution of IUD acceptors in family planning program by number of living children, age and type of area 1971, Taiwan area*

Characteristics	*Type of area*			
	Cities	*Urban townships*	*Rural townships*	*Total*
Number of living children				
0 or 1	10.3	6.1	4.1	6.8
2	24.8	18.9	15.0	19.5
3	29.3	28.5	27.4	28.3
4	20.0	23.3	24.8	22.7
5	9.0	12.7	14.9	12.2
6	4.0	6.1	8.0	6.1
7 or more	2.6	4.4	5.8	4.4
Total:				
Percentage	100.0	100.0	100.0	100.0
Number (a)	54,063	40,435	59,206	153,704
Age of acceptor				
Under 25	21.6	16.2	15.0	17.7
25-29	35.1	31.1	28.8	31.6
30-34	25.3	27.3	27.3	26.6
35-39	12.3	16.3	17.9	15.5
40 and above	5.7	9.1	11.0	8.6
Total:				
Percentage	100.0	100.0	100.0	100.0
Number (a)	54,582	40,715	59,609	154,906

(a) Excluding those whose age or number of children is unknown.
Source: Taiwan Provincial Committee on Family Planning, 1972.

on average about 3.5 months after a childbirth for nonlactating and 10.6 months for lactating loop acceptors (an overall average of 10.1 months) (Jain and Sun, 1972). If the women are recruited long after this period, there is a high possibility of their becoming pregnant; but for cases recruited before this period, there will be an overlap with postpartum amenorrhea. As Table 12.10 shows, loop acceptors in 1970 were recruited on average 22 months after childbirth, pill acceptors 31 months, and condom acceptors 37.5 months after. This period is much shorter for "spacers" than for "stoppers". The period also tends to be shorter among acceptors in the cities, those with higher education, those who are younger, and those who are accepting a method for the first time. The 1971 statistics show that the proportion of acceptors who accepted loops, pills or condoms within 6 months of a birth were 39.5 per cent, 21.2 per cent and 21.8 per cent, respectively. The 1970 proportion for pill and condom combined was 17 per cent; thus the overlap with postpartum amenorrhea seems to have increased over time as the program has put more emphasis on postpartum cases.

Table 12.10. *Mean length of interval between last childbirth and date of contraceptive acceptance by purpose of acceptance and method accepted, 1970 (in months)*

	Mean length of interval by method accepted		
Purpose of acceptance	*Loop*	*Pill*	*Condom*
For spacing	8.4	12.6	14.7
For stopping	25.7	35.5	42.9
Total	22.0	31.1	37.5

Source: Committee on Family Planning, 1970 and 1971.

III. USE-EFFECTIVENESS

CONTINUATION OF USE

The demographic impact of a family planning program also depends on how long the acceptors continue to use the accepted methods. As mentioned earlier, several island-wide follow-up surveys of IUD and pill acceptors have been carried out in Taiwan to measure continuation rates. The results have been reported in many places (Chow *et al.*, 1968; Potter *et al.*, 1969; Jain and Sun, 1970;

Table 12.11. *Observed and estimated continuation rates of IUD per 100 women for first segment and for all segments, Taiwan, 1968*

Months of use t	Retention rates per 100 IUD insertions							
	First segment					All segments		
	Observed	Estimation (1)		Estimation (2)		Observed	Estimation (1)	
	O	E_1	E_1-O	E_2	E_2-O	O′	E_1'	$E_1'-O'$
1	91.2	82.6	−8.6	92.9	+1.7	92.8	84.3	−8.5
2	85.7	80.3	−5.1	87.3	+1.6	88.5	82.7	−5.8
3	81.7	78.1	−3.6	82.6	+0.9	85.3	81.1	−4.2
4	77.4	76.0	−1.4	78.7	+1.3	82.1	79.6	−2.5
5	74.8	73.9	−0.9	75.4	+0.6	79.8	78.1	−1.7
6	72.3	71.9	−0.4	72.8	+0.5	77.5	76.6	−0.9
7	69.8	70.0	+0.2	70.0	+0.2	75.4	75.1	−0.3
8	67.5	68.0	+0.5	67.6	+0.1	73.5	73.7	+0.4
9	67.5	68.0	+0.5	67.5	−0.1	71.9	72.3	+0.4
10	63.7	64.4	+0.7	63.6	−0.1	70.1	70.9	+0.8
11	61.8	62.6	+0.8	61.7	−0.1	68.5	69.6	+1.1
12	59.8	60.9	+1.1	59.9	+0.1	66.9	68.3	+1.4
13	57.4	59.3	+1.9	58.3	+0.9	65.1	67.0	+1.9
14	55.9	57.8	+1.9	56.7	+0.8	63.8	65.7	+1.9
15	54.6	56.1	+1.5	55.1	+0.5	62.8	64.4	+1.6
16	53.2	54.6	+1.4	53.5	+0.3	61.6	63.2	+1.6
17	52.1	53.1	+1.0	52.1	0.0	60.7	62.0	+1.3
18	50.6	51.6	+1.0	50.8	+0.2	59.5	60.8	+1.3
19	49.5	50.2	+0.7	49.4	−0.1	58.3	59.7	+1.4
20	48.4	48.9	+0.5	48.1	−0.3	57.2	58.6	+1.4
21	47.4	47.5	+0.1	46.9	−0.5	56.3	57.4	+1.1
22	46.1	46.2	+0.1	45.6	−0.5	55.3	56.4	+1.0
23	44.5	45.0	+0.5	44.4	−0.1	54.1	55.3	+1.2
24	43.0	43.8	+0.8	43.3	+0.3	53.6	54.2	+1.2
25	41.5	42.6	+1.1	42.1	+0.6	52.0	53.2	+1.2
26	40.4	41.4	+1.0	41.0	+0.6	51.3	52.2	+0.9
27	39.3	40.3	+1.0	40.0	+0.7	50.1	51.2	+1.1
28	38.1	39.2	+1.1	38.9	+0.7	49.1	50.2	+1.1
29	37.1	38.2	+1.1	37.9	+0.8	48.2	49.3	+1.1
30	36.1	37.1	+1.0	36.9	+0.8	47.3	48.3	+1.0

Table 12.11. *Observed and estimated continuation rates of IUD per 100 women for first segment and for all segments, Taiwan, 1968 (continued)*

nths of se t	Retention rates per 100 IUD insertions							
	First segment					All segments		
	Observed	Estimation (1)		Estimation (2)		Observed	Estimation (1)	
	O	E_1	E_1-O	E_2	E_2-O	O'	E_1'	$E_1'-O'$
31	35.3	36.1	+0.8	35.9	+0.6	46.5	47.4	+0.9
32	34.4	35.1	+0.7	35.0	+0.6	45.4	46.5	+1.1
33	33.7	34.1	+0.4	34.1	+0.4	45.2	45.6	+0.4
34	32.9	33.2	+0.3	33.2	+0.3	44.3	44.8	+0.5
35	32.5	32.3	−0.2	32.3	+0.2	43.9	43.9	0.0
36	31.7	31.4	−0.3	31.4	−0.3	43.1	43.1	0.0
37	30.4	30.6	+0.2	30.6	+0.2	42.2	42.3	+0.1
38	29.7	29.7	0.0	29.8	+0.1	41.8	41.5	−0.3
39	29.4	28.9	−0.5	29.0	−0.4	41.3	40.7	−0.6
40	28.5	28.1	−0.4	28.3	−0.2	40.5	39.9	−0.6
41	27.6	27.4	−0.2	27.5	−0.1	39.3	39.4	−0.2
42	26.6	26.6	0.0	26.8	+0.2	38.5	38.4	−0.1
43	26.1	25.9	−0.2	26.1	0.0	37.9	37.7	−0.2
44	25.3	25.2	−0.1	25.4	+0.1	37.5	37.0	−0.5
45	24.9	24.5	−0.4	24.8	−0.1	37.1	36.3	−0.8
46	24.6	23.8	−0.8	24.1	−0.5	36.9	35.6	−1.3
47	23.9	23.2	−0.7	23.5	−0.4	36.0	34.9	−1.1
48	23.3	22.6	−0.7	22.8	−0.5	36.0	34.2	−1.8

Note: Estimation (1) A simple curve suggested by Mauldin and his colleagues (1968) was used:
$R_{(t)} = ae^{-rt}$
: $R_{(t)}$ = Retention rate at time t
a = Immediate expulsion rate
r = Monthly rate of termination

The values of a and r were estimated by the method of least squares:

For first segment	*For all segments*
$\hat{a} = 0.8486$	$\hat{a} = 0.8593$
$\hat{r} = 0.0276$	$\hat{r} = 0.0192$

Estimation (2) A formula suggested by Liu, Chow and Abbey (1971) is:
$R_{(t)} = 0.18e^{-0.3194\,t} + 0.82e^{-0.0266\,t}$
$R_{(t)}$ = Retention rate at time t

Source: Jain, A. K. and T. H. Sun, 1970. Liu *et al.*, 1972.

Avery, 1972a; Liu *et al.*, 1972). The important findings are summarized below:

1. *CONTINUATION OF IUD USE.* There has been criticism that "discontinuation rates for the IUDs inserted are so high as to mean little protection, little motivation for continued contraception and little reduction in fertility" (Potter *et al.*, 1968). Taiwan's experience has demonstrated, however, that as more data become available on the long-term use of the IUD, the continuation rates are not as bad as the estimates made at the early stages of the program based on a short period (two years or so) of experience. According to the third island-wide IUD follow-up survey (1968-1969), in which about 4,800 women were interviewed (a probability sample from the 435,000 who had had IUDs inserted during January 1964-June 1968), about 23.31 per cent were still using the original IUD at the end of 48 months ([5]). The continuation rate based only on the first segment of IUD use underestimates the period of protection provided by the IUD because it does not include the additional protection provided by reinsertions. About 30 per cent of those who had terminated the use of their first IUD had had a reinsertion by the time of this survey. If the reinserted IUDs are also taken into account, 36 per cent of the acceptors were still using an IUD four years after initial acceptance (Jain and Sun, 1970), as can be seen from Table 12.11. The average expected length of use of the first IUD is about 31 months, but when reinserted IUDs are also taken into account, the average length of IUD use is extended to about 45 months.

The continuation rates for all segments of IUD use calculated by Potter's life-table method tend to underestimate the number of current users because:

(1) the calculation excludes any other segments of IUD use after the one terminated by pregnancy;
(2) 35 per cent of those who terminated the first segment because of pregnancy had a reinsertion (the average reinsertion rate is 30 per cent).

Therefore, the number of loop users for the end of 1971 estimated using life-table rates is significantly smaller than that estimated from the other method in which IUD follow-up survey results are used directly. As R. Avery (1972a) noted, by examining the table prepared by Jain and Sun from the Taiwan IUD III survey in 1968, about

([5]) The results were based on Potter's life-table method (Potter, 1969).

50 per cent of the women who first accepted IUDs were using an IUD two years or more after the first insertion, irrespective of the period between first insertion and interview. After the first two years, reinsertions balance terminations so that the number of current IUD users remains constant (at about 50 per cent of the first acceptors) (see Table 12.21). Taking the proportion of acceptors still using an IUD at a specified time period after the first insertion, as shown in Table 12.21, adjusting for those above age 45 using Liu, Chow and Abbey's ratio, the number of current IUD (loop) users at the end of 1971 was estimated to be about 382,700, i.e., about 50 per cent of the women who had had first insertions or 21 per cent of the 1.8 million currently married women of child-bearing age (Table 12.12). This is 17.8 per cent higher than the 325,000 estimated using life-table rates.

Table 12.12. *Estimation of current IUD users at the end of 1971, applying the proportion of first IUD acceptors who are still using IUD at a specific time period after acceptance (obtained from Taiwan IUD III survey)*

The year when had first IUD insertion	*Number of first IUD insertions (1)*	*Percentage still using first IUD at the end of 1971 (2) (a)*	*Estimated number of women still using IUD at the end of 1971 (1) × (2)*
1964	49,496	49.0 (*b*)	24,253
1965	96,474	49.0 (*b*)	47,272
1966	105,680	49.0 (*b*)	51,783
1967	110,530	49.0 (*b*)	54,160
1968	101,533	49.0	49,751
1969	95,031	48.0	45,615
1970	99,024	54.0	53,473
1971	105,144	69.0	72,549
Total	762,912		398,856 (*c*)

(*a*) Derived from the percentage of women who had had an IUD inserted x years before the Taiwan IUD III survey interview (1968) and who were still using an IUD at the interview date. For example, the 1971 figure is the average proportion of those who had had a first IUD inserted 0-6 and 6-11 months previous to the interview and were still using an IUD at the time of interview (see Table 12.21).

(*b*) These percentages were estimated to be the same as 36-47 months, as rates after 24 months appeared to be constant.

(*c*) If this figure is adjusted by the Liu and Chow model 2 to eliminate those who were over 45 years old at the time of interview, the continued loop users at the end of 1971 would be:

$$398{,}856 \times 0.9595 = 382{,}702$$

Source: R. Avery, 1972b. (Corrections were made on 1970 and 1971 first insertion figures.)

Using the same method, the number of women currently using an IUD at the end of other years can be estimated as in Table 12.13. The proportion of current IUD users among the cumulative total of

IUD insertions decreases from 65.2 percent in 1964 to 40.9 per cent in 1971, and the proportion of current IUD users among the cumulative total of first IUD insertions from 66.3 per cent to 50.2 per cent over the same period. The rate of decrease has been very low since 1967 (Figure 12.2). This implies that the current number of IUD users will be maintained at about half the total first insertions ever made. The proportion of currently married women aged 20-44 using IUDs had increased to 21.1 per cent in 1971, though the increase in this rate has been slower in recent years than for the years immediately after the program was initiated. This estimate seems to fit the actual situation quite well. The proportion of married women aged 20-44 who were currently using IUDs at the end of the KAP III survey (January 1970) was 16.1 per cent and the estimated proportion of currently married women aged 20-44 using IUDs at the end of 1969 was 16.5 per cent.

Table 12.13. *"Prevalence" of IUD use in Taiwan at the end of specified year, 1964-1971*

End of year	*No. of married women 20-44 (1,000s)*	*Cumulative total no. of IUD insertions (1,000s)*	*No. of women currently using an IUD (a) (1,000s)*	*% of current IUD users among*		
				Total no. of married women	*Total no. of IUD insertions*	*Cum. no. of first IUD insertions*
1964	1,555	50.3	32.8	2.1	65.2	66.3
1965	1,594	149.5	89.5	5.6	59.9	61.3
1966	1,622	260.7	142.7	8.8	54.8	56.7
1967	1,648	381.8	195.6	11.9	51.2	54.0
1968	1,687	505,5	241.8	14.3	47.8	52.2
1969	1,728	635.8	284.7	16.5	44.8	51.0
1970	1,769	779.1	331.8	18.8	42.6	50.4
1971	1,811	934.7	382.7	21.1	40.9	50.2

(a) Estimated using the proportion of women who were still using an IUD at the time of Taiwan IUD III Survey (1968). For Methodology, see Table 12.12.

The follow-up studies of 5,832 women who accepted IUDs in Taichung City from 1962 to 1965 provide long-term observation of IUD use under a clinical follow-up situation. The results of the fourth survey carried out at the end of 1971 show that 33.6 per cent of the IUD acceptors were still using the original IUD at the end of 48 months and 22 per cent at the end of 96 months (Table 12.14), much higher rates than those observed in the island-wide survey results. When

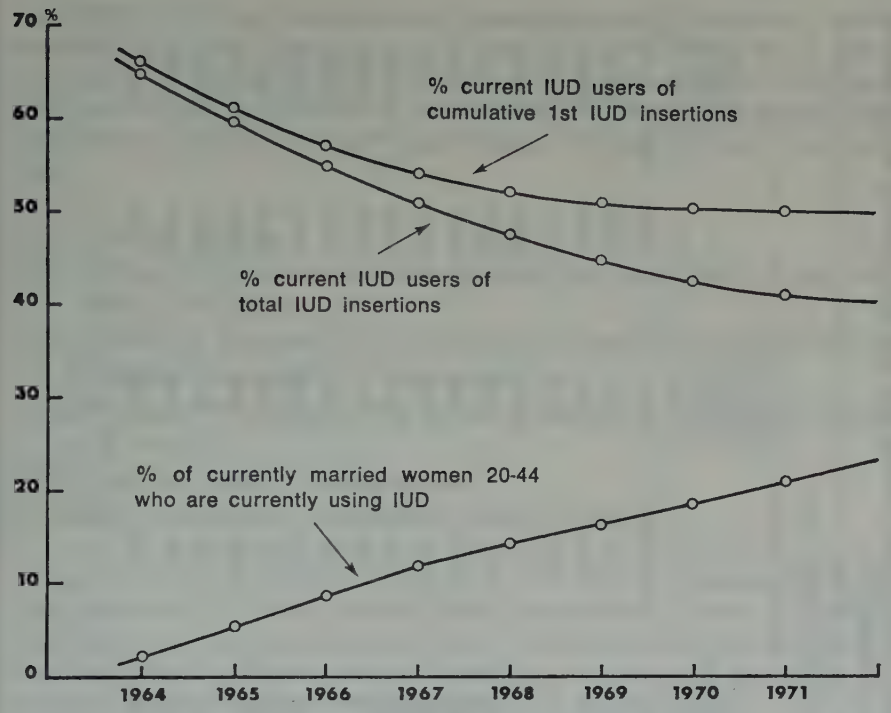


Fig. 12.2. *Prevalence of IUD use in Taiwan area*

reinserted IUDs were taken into account, the proportion still using IUDs increased to 52.5 per cent at the end of 48 months and 44.9 per cent at the end of 96 months (Moots and Chen, 1973). These are remarkably high continuation rates.

Avery (1972a) pointed out that the "six months conditional termination rates" for all causes drop dramatically from 23 per cent of all users during the first six months to only 4 per cent of continuing users between 78 and 84 months; thus, women who do not drop out in the first two years are likely to retain their original IUD for a long period thereafter. This means that estimates such as the numbers of future IUD users and average life of a device, which are based on one or two years of experience and assume constant termination rates (Mauldin, 1968), will be underestimates (Jain and Sun, 1970). The Liu, Chow and Abbey (1972) model, which assumes a decreasing rate of termination, seems to fit the observed rates rather well (Table 12.11).

It has also been reported that IUD continuation rates are strongly

Table 12.14. *Current pregnancy or contraceptive status of women at six months intervals after first IUD insertion, fourth Taichung IUD follow-up survey, 1971*

No. of months after first insertion	Pregnancy or contraceptive status (%)							Number of women involved (a)
	Original IUD	Reinserted IUD	Sterilized	Other contraceptive methods	Pregnant	Not using any method and not pregnant	Total	
6	77.6	4.5	0.2	3.9	5.4	8.3	100.0	5,672
12	65.0	6.7	1.1	6.4	9.7	11.1	100.0	5,596
18	56.0	9.0	2.3	7.9	9.8	14.9	100.0	5,533
24	48.7	11.1	3.7	10.0	9.0	17.6	100.0	5,456
30	43.6	13.8	5.2	10.8	7.8	18.7	100.0	5,362
36	38.9	16.1	6.4	12.6	6.3	19.6	100.0	5,309
42	35.8	17.9	7.9	13.3	5.8	19.2	100.0	5,225
48	33.6	19.3	9.2	14.7	4.9	18.4	100.0	5,168
54	31.9	20.5	10.3	14.6	3.9	18.8	100.0	5,044
60	29.7	21.3	11.4	15.1	3.2	19.2	100.0	4,997
66	28.3	22.1	12.4	14.9	2.7	19.6	100.0	4,887
72	26.7	22.8	13.4	14.8	2.3	19.9	100.0	4,827
78	25.2	23.8	15.2	14.3	1.9	19.5	100.0	4,496
84	23.9	23.8	16.2	14.7	1.5	19.9	100.0	3,836
90	22.5	23.2	18.2	14.4	1.4	20.3	100.0	2,823
96	22.0	22.9	18.6	14.0	1.4	21.2	100.0	2,033

(a) Women lost to follow-up, menopausal, dead, widowed, separated, or divorced are not included in the total.
Source: Moots, 1972.

correlated with the characteristics of acceptors (Chow *et al.*, 1968; Potter *et al.*, 1969; Freedman and Takeshita, 1969; Peng *et al.*, 1970; Hermalin and Chow, 1971). In the third island-wide IUD follow-up survey it was observed that women with the following characteristics had higher termination rates than others at 36 months after first insertion:

(1) younger women;
(2) women with fewer previous pregnancies, live births, living children or living sons;
(3) those who accepted an IUD for spacing their children rather than limiting the growth of their families;
(4) those who had used other forms of contraception before accepting an IUD;
(5) those who live in urban areas;
(6) those with better education (Table 12.15).

Avery (1972a) observed, however, in the fourth Taichung IUD follow-up survey (1971) data, that these correlations appear to diminish after the first two years of IUD use. He asserted that the effect of background characteristics of acceptors on termination rates can then be ignored and that acceptors can be treated as a homogeneous group as far as these characteristics are concerned (6).

2. *CONTINUATION OF PILL USE.* The estimation of continuation rates is much more difficult for the pill than for the IUD because of the high irregularity of pill use. An island-wide pill acceptor follow-up survey was carried out in July-August 1968 to measure the use-effectiveness of the pill, which was introduced into the program in January 1967 (Chen *et al.*, 1970). The continuation rates, calculated by the life table method, are much poorer than those for IUDs. Of the 2,217 respondents, 131 (5.9 per cent) never took the first pill. For all segments, i.e., the period from the time of acceptance to the last interruption of pill use, 24 per cent of the pill acceptors had stopped taking the pills at the end of the first month, 36 per cent at the end of the third month, 51 per cent at the end of the sixth month and 69 per cent at the end of one year. Side effects were mentioned as the most important reason (69 per cent) for terminating pill use. This was probably related to the following:

(6) Observation of the tables presented by him, however, seem to indicate that the differentials persist through the 24-48 month interval.

Table 12.15. *Cumulative gross rates per 100 first insertion*

Characteristics at time of first insertion	*Number*	*Type of termination*			
		Pregnancy	*Expulsion*	*Removal*	*Total*
1. *Age of wife*					
24 or less	611 [a]	19.1 (3.1) [b]	21.3 (2.8)	77.4 (4.9)	85.6 (1.9)
25-29	1,545	17.3 (1.7)	20.3 (1.6)	62.6 (2.6)	75.3 (1.3)
30-34	1,435	15.3 (1.5)	13.8 (1.3)	52.4 (2.4)	65.2 (1.5)
35-39	915	10.1 (1.5)	6.6 (1.1)	47.4 (2.7)	55.8 (1.9)
40 or more	292	5.3 (1.9)	5.7 (1.7)	47.5 (4.8)	53.1 (3.4)
2. *Number of pregnancies*					
2 or less	492 [a]	19.0 (3.5)	25.5 (3.5)	82.2 (5.6)	89.3 (1.9)
3	766	14.9 (2.3)	22.1 (2.4)	65.4 (3.9)	77.0 (1.9)
4	910	15.7 (2.0)	17.2 (1.8)	57.8 (3.2)	70.5 (1.8)
5 or more	2,605	13.0 (1.0)	10.2 (0.8)	50.5 (1.7)	61.3 (1.1)
3. *Number of live births*					
2 or less	624 [a]	16.9 (2.9)	23.8 (3.0)	82.9 (4.9)	89.2 (1.7)
3	986	16.5 (2.1)	20.0 (1.9)	62.2 (3.3)	74.7 (1.7)
4	1,072	15.9 (1.9)	14.9 (1.5)	55.0 (2.8)	67.8 (1.7)
5 or more	2,112	12.2 (1.1)	10.0 (0.9)	49.3 (1.9)	59.9 (1.2)
4. *Number of living children*					
2 or less	695 [a]	19.4 (3.0)	23.5 (3.0)	81.1 (4.6)	88.3 (1.6)
3	1,133	15.5 (1.9)	19.5 (1.7)	60.3 (3.0)	73.0 (1.6)
4	1,167	15.4 (1.8)	14.8 (1.5)	54.8 (2.7)	67.4 (1.6)
5 or more	1,801	11.9 (1.2)	9.1 (0.9)	48.6 (2.0)	58.8 (1.4)
5. *Number of living sons*					
1 or less	1,216	16.7 (2.1)	17.3 (1.6)	70.1 (3.1)	79.4 (1.5)
2	1,974	14.0 (1.2)	16.2 (1.2)	55.3 (2.1)	67.8 (1.3)
3	1,017	14.7 (1.8)	11.8 (1.3)	49.6 (2.7)	62.1 (1.8)
4 or more	591	11.3 (1.9)	8.4 (1.6)	49.3 (3.3)	58.8 (2.3)
6. *Number of induced abortions*					
None	3,619	12.9 (0.9)	15.8 (0.9)	56.9 (1.6)	68.4 (0.9)
One	707	17.7 (2.2)	11.5 (1.8)	53.9 (3.3)	66.4 (2.0)
Two or more	472	18.2 (2.9)	8.5 (1.6)	58.4 (4.1)	68.8 (2.4)

standard errors 36 months after insertion, Taiwan, 1968

haracteristics at time of first insertion	Number	Type of termination			
		Pregnancy	Expulsion	Removal	Total
Number of other fetal deaths					
None	3,807	15.2 (1.0)	15.2 (0.8)	55.9 (1.5)	68.3 (0.9)
One or more	991	11.2 (1.6)	11.7 (1.5)	58.5 (2.9)	67.5 (1.7)
IUD for limiting or spacing					
Limiting	4,031	13.7 (0.9)	13.3 (0.7)	52.4 (1.4)	64.4 (0.9)
Spacing	760 (a)	20.0 (3.5)	22.9 (2.8)	82.9 (4.3)	89.4 (1.4)
Use of contraception					
Did not use	3,797	13.9 (0.9)	15.1 (0.8)	56.8 (1.5)	68.4 (0.9)
Ota ring	492	14.7 (2.4)	11.7 (2.1)	49.0 (3.8)	61.6 (2.5)
Other method	505	16.6 (2.6)	11.7 (2.0)	62.1 (4.2)	72.1 (2.3)
Cost of insertion					
Cost	3,553	14.7 (0.9)	14.3 (0.8)	55.3 (1.5)	67.3 (0.9)
Free	1,060 (a)	13.1 (2.2)	13.7 (1.6)	62.0 (3.7)	71.5 (2.2)
Type of device					
Loop I	376	20.3 (3.0)	11.9 (2.1)	56.1 (4.5)	69.1 (2.5)
Loop II-IV	2,463	12.0 (1.2)	14.2 (1.0)	56.7 (2.0)	67.3 (1.2)
Unknown	1,959	16.0 (1.3)	15.4 (1.2)	56.7 (2.1)	69.2 (1.2)
Place of residence					
Large city	1,008	19.3 (2.1)	14.4 (1.7)	70.3 (3.3)	79.5 (1.5)
Small city	361	16.7 (3.2)	9.3 (1.9)	58.0 (4.7)	68.2 (2.8)
Small town	1,426	13.6 (1.5)	15.6 (1.4)	53.3 (2.4)	66.0 (1.5)
Rural	2,003	12.1 (1.2)	14.6 (1.1)	51.6 (2.0)	63.7 (1.3)
Education of wife					
No formal or tutor	2,432	11.7 (1.0)	14.0 (1.0)	55.8 (1.9)	66.4 (1.2)
Primary graduated	1,808	16.2 (1.4)	15.3 (1.2)	54.5 (2.1)	67.7 (1.3)
Junior or above	557	20.8 (3.2)	13.8 (2.1)	66.7 (4.4)	77.3 (2.1)
l sample	4,791	14.4 (0.8)	14.5 (0.7)	56.8 (1.4)	68.3 (0.8)

(a) Fewer than 50 women were exposed at 36 months.
(b) Standard errors are shown in parentheses.
Source: Jain and Sun,1970, pp. 17-18.

(1) the pill introduced was one which contained a high dosage of estrogen;
(2) the United States Congress hearings on pill use created many negative rumors about possible side effects, so that women lost confidence in the pill; and
(3) factors related to pill supply and personal reasons.

A lower estrogen dosage pill was introduced in 1968 to replace the first pill used and another survey is planned to determine whether the continuation rate has increased.

Another method is being used to measure the proportion of all acceptors who are continuing use. Theoretically, the number of pill cycles distributed to acceptors each month should be roughly equal to the number of continuing users. One cycle of pills is provided at acceptance and three cycles on each subsequent attendance. The difference should balance out in the long run. Therefore, the pill continuation rate at month (t) could be estimated by:

$$R(t) = \frac{\text{No. of cycles distributed in months t to previous acceptors}}{\text{Cumulative no. of new acceptors up to months } (t-1)}$$

According to this estimate, the proportion continuing pill use dropped from 100 per cent in early 1967 to 40 per cent by mid-1968, to 30 per cent in mid-1969, to 25 per cent in mid-1971, and to 20 per cent towards the end of 1972. Of course, the actual number of current users has increased from 20,000 in 1968 to 60,000 in 1972, based on the number of cycles distributed monthly.

3. *CONTINUATION OF CONDOM USE.* The use of condoms is even more irregular than that of the pill. An island-wide condom acceptor follow-up survey is still in progress, and a "life table" of continuation rates is not yet available. Using the method stated in the pill section, it is possible, however, to estimate roughly the proportion of total acceptors who continue to use the condom in a specific month. If it is assumed that a dozen condoms will be used for about 1.5 months, then the condom continuation rate at month (t) could be calculated by:

$$R(t) = \frac{(\text{Dozens of condom distributed in month t to previous acceptors}) \times 1.5}{\text{Cumulative no. of new acceptors up to month } (t-1)}$$

According to this estimation process, the rate has been maintained at between 20 and 30 per cent during 1970 and 1971 and was still

above 15 per cent at the end of 1972. The number of current users has increased from 10,000 in 1970 to 25,000 in 1972.

BIRTHS AVERTED BY THE USE OF PROGRAM METHODS

Another important question in estimating the effect on fertility of a family planning program is "how many births have been averted as a result of a national family planning program?" Cernada (1970) deals specifically with the case of Taiwan in the "Program Effect on Fertility" section of the *Taiwan Family Planning Reader,* which includes articles dealing with fertility limitation after IUD insertion (Freedman *et al.,* 1971), a matching study of loop and non-loop acceptors to estimate demographic impact of the family planning program (Chang *et al.,* 1969) and the use of areal analysis to isolate the impact of the family planning program from social and economic factors (Hermalin, 1971).

Mauldin (1968) listed five major approaches to this question:

(1) comparison of the fertility trend before and after the initiation of the program to measure the difference in birth rates attributable to the program;
(2) comparison of the fertility of acceptors before and after the acceptance of program methods;
(3) comparison of the expected number of births without contraception and induced abortion with the actual number of births;
(4) comparison of fertility between acceptors and non-acceptors with comparable characteristics;
(5) calculation of the number of births which could be prevented by use of a program method, taking into account the fertility level of acceptors and the couple-years of effective contraception.

The application of the third and the last approaches will be discussed in this section, while the first, second and fourth approaches will be discussed in the following sections.

1. *BIRTHS AVERTED BY THE USE OF IUDs.* Chow (in Ross, 1966) estimated that five IUD insertions in Taiwan would avert one birth per year for at least five years, starting the year following the insertion, or that each IUD inserted would prevent on average one birth within five years because (according to Ross):

"(1) of five IUDs inserted this year, two will come out before they have time to do much good. The others will stay in at least five years;

(2) the marital fertility rate among acceptors would be about 333 per thousand in the absence of a program."

Following the second island-wide IUD follow-up survey (1966), this rough estimate was revised (Chow, 1968) to indicate that "four loop insertions prevent one live birth in the year following insertion." In this case the number of births averted was estimated as the difference between the fertility of IUD acceptors after acceptance (excluding the first nine months) and their expected fertility without the use of the IUD. The latter was estimated by adjusting the fertility of acceptors before acceptance (a three year average) for the rate of marital fertility change in Taiwan between 1964 and 1965. The result was adjusted 4 per cent downwards to account for fertility reduction due to aging. This method overestimates the effect of the IUD in that the measurement of the fertility of acceptors after acceptance was based on the number of live births, while it is known that more than half (62 per cent) of the pregnancies of acceptors after acceptance of an IUD (regardless of use status) were terminated by induced abortions. Thus the estimate includes not only the effect of the use of the IUD but also the effect of induced abortion. The estimate does not show how many births an IUD insertion can prevent in its life span, so it has not been used here to calculate the total number of births averted by all the IUDs inserted.

Potter (1969) has made a careful estimate of the number of births averted by each initial IUD insertion for women of different age groups, using Taichung IUD follow-up survey data and taking three alternative fertility assumptions (Table 12.16).

Table 12.16. *Births averted per first segment of IUD use by age of acceptor and assumptions in respect of potential fertility*

Assumption concerning fertility in absence of IUD	*Births averted per first segment of IUD*				
	Age group				*All ages*
	20-24	*25-29*	*30-34*	*35-39*	
Maximal (no family planning)	0.54	0.80	1.12	1.10	0.94
Medium	0.54	0.68	0.72	0.54	0.64
Conservative (everybody substituting)	0.47	0.51	0.46	0.24	0.43

Source: Potter, 1969.

Judging from the characteristics of acceptors stated in the previous section, the actual fertility situation in Taiwan probably lies somewhere between "medium" and "maximal" with a trend moving closer to the "medium" assumption (Table 12.4, 4th panel). Let us assume that:

(1) the "assumption concerning fertility in the absence of the IUD" can be represented by the proportion of women who have never used contraceptive methods before acceptance; i.e., the assumption "maximal" is equivalent to the statement "100 per cent of acceptors had never used contraception before acceptance" and the assumption "medium" to "50 per cent of acceptors had no previous use of contraception" shown in Table 12.4 ([7]).
(2) the numbers of births averted per first segment of IUD use for each age group are linearly distributed between the two fertility assumptions—maximal and medium.

With these two assumptions, the number of births averted per first segment of IUD use, corresponding to the proportions that never used contraception before acceptance shown in Table 12.4, can be interpolated and the total number of births averted by the first segment of IUD insertion in Taiwan from 1964 to 1971 can be calculated as in Table 12.17. About 542,000 births could have been averted by 700,000 first IUD insertions between 1964 and 1971 among women aged 20-39, an average of .78 births per first segment of IUD use. The number of births averted by one year's first IUD insertions lies between 63,000 and 83,000 (except for 1964), equivalent to about 16-18 per cent of the 400,000 or so annual live births. Even though the number of births averted by one year's IUD insertions, shown in Table 12.17, will be spread out over several years following the year of insertion, the fairly constant number of first insertions done every year seems to indicate that, in the long run, the number of births that could be averted each year by the first IUD insertions will be maintained at about 60,000-80,000 starting around 1968, assuming that the charac-

([7]) Actually, the contraceptive status before acceptance is not equivalent to that after acceptance, even in the absence of an IUD, because the proportion practicing contraception will be higher in the later years. However, this deficiency could be overcome to some extent by the fact that "the proportion who never used contraceptives before acceptance", shown in Table 12.4, is too low because it refers to all acceptors including those who had accepted an IUD before. If the proportion is limited to the initial acceptors of IUDs, it would be much higher. This assumption is valid only when the proportion of initial IUD acceptors who will accept some form of family limitation measures in the absence of an IUD will be roughly equal to the proportion who had an IUD reinserted.

Table 12.17 *Number of births averted by first segment of IUD use for acceptors aged 20-39, 1964-1971, Taiwan area*

Age group	Year of acceptance								Total
	1964	1965	1966	1967	1968	1969	1970	1971	
					Number of first IUD acceptors				
20-44	3,131	7,046	9,139	10,375	11,589	14,081	17,471	22,247	95,079
25-29	12,733	25,873	30,347	33,202	32,518	32,572	34,439	36,986	238,670
30-34	17,031	32,071	33,829	35,139	30,614	26,331	25,998	25,520	226,533
35-39	12,140	22,293	22,415	21,729	17,882	14,277	12,231	12,761	135,728
Total	45,035	87,283	95,730	100,445	92,603	87,261	90,139	97,514	696,010
					Number of births averted per first segment use (a)				
20-24	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
25-29	0.76	0.76	0.76	0.75	0.73	0.72	0.71	0.70	0.73
30-34	0.98	0.98	0.98	0.95	0.89	0.83	0.81	0.79	0.90
35-39	0.91	0.91	0.91	0.86	0.78	0.70	0.67	0.64	0.82
Total (b)	0.87	0.86	0.85	0.82	0.77	0.72	0.70	0.68	0.78
					Total number of births averted				
20-24	1,691	3,805	4,935	5,603	6,258	7,604	9,434	12,013	51,343
25-29	9,677	19,663	23,064	24,902	23,738	23,452	24,452	25,890	174,838
30-34	16,690	31,430	33,152	33,382	27,246	21,855	21,058	20,161	204,974
35-39	11,047	20,287	20,398	18,687	13,948	9,994	8,195	8,167	110,723
Total	39,105	75,185	81,549	82,574	71,190	62,905	63,139	66,231	541,878

(a) Estimated by applying Potter's estimate of the number of births averted by each initial IUD insertion (Table 12.16 and the proportion of IUD acceptors who have never used contraception before acceptance of IUD, assuming that Potter's "maximum" effect applies to 100 per cent never used before and "medium" applies to 50 per cent of previous users). (See text for detailed discussion.)

(b) Weighted average.

Source: (for number of IUD acceptors). Taiwan Provincial Committee on Family Planning, 1972.

teristics of first IUD acceptors do not change ([8]). This is about 17 per cent of the annual number of births in recent years, which has been maintained at about 380,000-390,000. Actually the effect of the IUD program should be much greater than the estimate made above because the estimation does not include first insertions to women younger than age 20 or older than age 40, which form about 7-8 per cent of the total, nor the effect of reinsertions, which have been increasing during recent years (they formed about 35 per cent of all insertions in 1971). Unfortunately, the exact effect of a reinserted IUD has not been calculated. It should be noted that Potter's calculation was based on Taichung clinic data, which show a higher continuation rate than the island-wide sample. For example, the mean life span of the first segment was estimated as 38.2 months in Taichung while the corresponding figure was only 31 months for the island-wide sample, a figure which increases to 45 months when reinsertions are taken into account. Therefore, it can be said that Potter's estimate lies somewhere between the effects of the first segment for Taichung and of all segments for the island-wide sample. Recalculation should be made based on the island-wide sample when data become available.

The total number of births averted by the IUD program can also be estimated roughly from the number of women who are currently using the IUD at a specific time. If we assume that:

(1) the number of women using an IUD throughout a year can be represented by the current number of IUD users at the mid-year point;
(2) 90 per cent of the current users will use it effectively and prevent births which will be shown in the following year;
(3) the fertility of IUD users in the absence of an IUD is 328 per thousand, which is the difference between fertility before accepting the IUD (the 3 year average is 403 per thousand from the third Taiwan IUD follow-up survey) and the fertility of acceptors in the second year after acceptance (75 per thousand),

then the number of births averted by the IUD program in a specific year can be calculated as in Table 12.18. On this basis, the IUD

([8]) Actually, the characteristics of first IUD acceptors have been changing as was discussed in the second section of this chapter. The influences of the changes are in two directions: since the more recent acceptors are younger, their fertility will be higher; but more of them are "spacers", a factor resulting in lower use-effectiveness. It is very likely that the net effect of the changes will result in fewer births prevented by first IUD insertions later in the program.

Table 12.18. *Number of births averted by the program, by method and by year, 1965-1972, Taiwan area*

Year	Loop		Pill		Condom		Total
	Number of current users (mid-yr.) [a] (1)	Number of births averted [b] (2)	Number of cycles distributed (3)	Number of births averted [c] (4)	Number of dozens distributed (5)	Number of births averted [d] (6)	Number of births averted (2)+(4)+(6)
1964	16,400	—	—	—	—	—	—
1965	61,150	4,841	—	—	—	—	4,841
1966	116,100	18,051	—	—	—	—	18,051
1967	169,150	34,273	109,159	—	—	—	34,273
1968	218,700	49,933	230,660	1,928	—	—	51,861
1969	263,250	64,560	298,690	4,074	—	—	68,634
1970	308,250	77,711	405,781	5,275	119,960	—	82,986
1971	357,250	90,995	671,042	7,167	251,797	2,951	101,113
1972	—	105,460	—	11,852	—	6,194	123,506
Total		445,824	1,715,332	30,296	371,757	9,145	485,265

(*a*) Estimated from Table 12.13.

(*b*) (Number of current users in the previous years) × 0.9 × 0.328, where 0.9 is the effectiveness of IUD use; 0.328 is the fertility of users in the absence of an IUD, estimated from the difference between fertility before insertion (3 years average = 403) and the fertility of acceptors in the second year after acceptance (75).

(c) (Number of pill cycles distributed in the previous year) ÷ 13 × 0.7 × 0.328, where 0.7 is the effectiveness of use of pill, i.e., 70 per cent of the pills distributed were used effectively; the figure is divided by 13 to calculate the couple years of protection; and 0.328 is the fertility of pill acceptors in the absence of the pill.

(*d*) (Number of dozens of condoms distributed in previous year) ÷ 8 × 0.6 × 0.328, where 0.6 is the effectiveness of condom use, i.e., 60 per cent of the condoms distributed were used effectively; the figure is divided by 8 because it is assumed that two dozen condoms are enough for three months' use; and 0.328 is the fertility of condom acceptors in the absence of condoms.

Source: (for number of pill cycles and condoms distributed). Taiwan Provincial Family Planning Committee, 1972.

program in Taiwan had averted about 446,000 births by the end of 1972. This figure should be adjusted 4 per cent downwards, as suggested by Chow (1968) to account for the effect of aging on the fertility of acceptors. The method ignores the increasing possibility of use of other contraceptive methods in the absence of the IUD.

2. *BIRTHS AVERTED BY USE OF THE PILL AND CONDOM.* The number of births averted by the use of pills and condoms can be estimated roughly by using the concept of couple-years of protection.

Let us assume, based on past observation, that:

(1) 70 per cent of the pill cycles distributed in a year will be used effectively to prevent births in the following year;

e 12.19. *Number of births prevented inside and outside the family planning program, 1965-1972, Taiwan area*

r	Expected number of births without fertility control (a)	Actual number of births	Number of births prevented			Percentage of births prevented inside the program
			Total	Outside the program	Inside the program	
5	550,033	406,604	143,429	138,588	4,841	3.3
6	559,284	400,714 (b)	158,570	140,519	18,051	11.4
7	568,314	388,676 (b)	179,638	145,365	34,273	19.1
8	577,858	394,260	183,598	131,737	51,861	28.2
9	589,623	390,728	198,895	130,261	68,634	34.5
0	603,583	394,015	209,568	126,582	82,986	39.6
1	616,218	380,424	235,794	134,681	101,113	42.9
2	629,000 (c)	365,749	263,251	139,745	123,506	46.9
l			1,573,492	1,088,227	485,265	30.8

(a) Estimated using the following age-specific marital fertility rates:

Age of women	Fertility rate (per 1,000)
15-19	500
20-24	490
25-29	420
30-34	330
35-39	230
40-44	130
45-49	36

(b) Adjusted for the effect of the 1966 census.
(c) Estimated figure.

Source: Table 12.18 and Taiwan Demographic Fact Book, Ministry of the Interior, 1972.

(2) use of 13 cycles of pills is equivalent to one couple-year of protection;
(3) the fertility of pill users in the absence of the pill is 328 per thousand, as was assumed in the case of IUDs above (9).

With these assumptions, the number of births averted by pills distributed in the program can be estimated as in Table 12.18. By 1972, about 30,000 births had been averted by the distribution of about 1.5 million cycles of pills during 1967-1971. This is roughly 57 cycles of pills per birth prevented.

In the case of condom use, let us assume, also based on past observation, that:

(1) 60 per cent of the condoms distributed in a year will be used effectively to prevent births in the following year;
(2) use of 8 dozen condoms is equivalent to one couple-year of protection;
(3) the fertility of condom-users in the absence of condoms is also 328 per thousand, as mentioned above (9).

As shown in Table 12.18, 370,000 condoms distributed in 1970 and 1971 should have prevented about 9,000 births in 1971 and 1972.

3. *NUMBER OF BIRTHS AVERTED BY THE PROGRAM.* The last column of Table 12.18 shows the estimated total number of births averted by the use of the three program methods in each year. It makes it possible to estimate the extent of the contribution of the family planning program in reducing the birth rate.

By applying a set of fertility rates reached in the absence of fertility control practice, a theoretical maximum number of births (i.e., an expected number of births without contraception and induced abortion) can be calculated. The difference between this theoretical number of births and the actual number of births can be regarded as the number of births averted by the practice of fertility control methods both inside and outside the program. Since the number of births prevented by the program methods has been estimated as above (Table 12.18), it is possible to show the share of the program in preventing births each year. Table 12.18 shows the results of this calculation with the assumption that the age-specific marital fertility in the absence of fertility control practice is as follows:

(9) The assumption of fertility of pill and condom users may not be valid because of the difference in characteristics among users of the three methods. This assumption was adopted due to lack of better knowledge.

Age of women	*Fertility rate* (per 1,000)
15-19	500
20-24	490
25-29	420
30-34	330
35-39	230
40-44	130
45-49	36

These results indicated that the births prevented by couples using program methods increased from 11.4 per cent of all births prevented in 1965 to 46.8 per cent in 1972. The number of births prevented by couples using non-program methods reached 145,000 in 1967 but decreased to 127,000 in 1970, probably due to the substitution effect of program methods. It increased again to 140,000 in 1972. The proportion of births prevented inside the program in 1970 was 39.6 per cent which is fairly close to the result of the third island-wide KAP survey (in early 1970) which showed that about 38 per cent of the women who were currently using some form of contraception were using program methods (IUD and pills).

ASSESSMENT OF THE DEMOGRAPHIC IMPACT OF THE IUD PROGRAM THROUGH MATCHING STUDY

The effect of a program can also be measured by a fertility comparison between acceptors and non-acceptors with similar characteristics.

Chang, Liu and Chow (1969) carried out a study in four townships near Taichung City to measure the effect of the IUD program. This study takes advantage of the good household record-keeping system. A total of 3,181 IUD acceptors in 1964, 1965 and 1966 were selected from the sample area and an equal number of "matched cases" from the non-acceptors in the townships using the following criteria: age at time of insertion, education, open interval at the time of insertion and number of live births. All the acceptors were assumed to have accepted the IUD on July 1 of each year, and the cut-off date for observation was December 31, 1967. The average periods of observation were 39.2 months, 31.2 months and 18.4 months for 1964, 1965 and 1966 acceptors, respectively.

The main conclusions drawn from the study were as follows:

(1) The fertility of IUD acceptors before acceptance was 54 per cent

higher than the average fertility of married women in the townships.

(2) Acceptors in the IUD program had fertility rates of 381 per thousand before and 77 per thousand after initial entry into the program—an 80 per cent reduction. The corresponding fertility decline for the "matched cases" was 48 per cent, from 377 per thousand before to 195 per thousand after the reduction date (Table 12.20). Therefore the net effect ([10]) of the IUD program on fertility was a 32 per cent reduction. The amount of the decline may have been exaggerated by the fact that those who become acceptors do so after a recent birth which inflates pre-insertion fertility and also that for at least a time afterwards lower fertility is to be expected. These effects, however, were reduced in the study by taking into account the open birth interval as one of the matching criteria.

(3) On average, 159 live births would be prevented per 1,000 insertions during the first year, beginning nine months after the first insertion. The number of births averted would decrease to 113 during the second year and 99 during the third year.

(4) The impact of the IUD is significantly greater among women of lower educational status, those most in need of a family planning program.

(5) The impact of the IUD program goes beyond termination of the IUD.

This study could have been followed by an interview survey to find out about the contraceptive history of both the IUD acceptors and the matched cases. The results of this type of study would provide clues to understanding more thoroughly the impact of the IUD program, including its substitution effect.

IV. ANALYSIS OF FERTILITY

FERTILITY AFTER INSERTION OF AN IUD OR ACCEPTANCE OF THE PILL

1. *IUD.* An important feature of a family planning program is that its effect may go beyond the use-effectiveness of a particular contraceptive method provided in the program. The total effect of the program on the fertility of acceptors should be one of the most important

([10]) This is not the net effect of the IUD itself but the effect of the IUD program, including the effect of fertility control measures taken after termination of the IUD.

Table 12.20. *Fertility before and after first acceptance of IUD, fertility decline after acceptance and births averted, by year of first acceptance, matching study area, Taiwan, 1968*

Year of first acceptance	Sample size	Fertility rate per 1,000				Fertility decline (%)			Births averted per 1,000
		Before		After					
		Acceptors	Matches	Acceptors (a)	Matches	Acceptors	Matches	Index	
1964	463	384	390	70	172	82	56	58	99
1965	1,434	375	370	82	193	78	48	58	113
1966	1,284	387	377	74	227	81	40	68	158
Total	3,181	381	376	77	195	80	48	61	121

(a) Nine months were deducted for each acceptor in the calculation of post-insertion fertility because the acceptors were not pregnant when the IUDs were inserted. They could not have a live birth within the following nine months. The same reduction was made for the "matches" because they were also selected in such a way that they could not have a live birth within nine months.

Source: Chang *et al.*, 1969.

Table 12.21. *Current family planning status for IUD acceptors by number of months since first insertion, Taiwan, 1968*

Current family planning status	Months between first insertion and interview									
	0-5	6-11	12-17	18-23	24-29	30-35	36-41	42-47	48+	Total
	%	%	%	%	%	%	%	%	%	%
IUD still in place	76	62	58	50	46	51	48	50	41	51
Currently using IUD or other contraception(a)	85	76	76	72	70	71	69	72	71	72
Currently using IUD or other contraception, or aborting all pregnancies	85	78	78	76	76	76	75	78	76	77
Currently using IUD or other contraception, aborting all pregnancies or having none	100	94	91	87	87	89	87	90	88	90
Unprotected: At least one live birth and not currently using contraception	—	1	3	6	7	9	12	9	11	7
Unprotected: Currently pregnant (only one pregnancy)	—	5	6	7	6	2	1	1	1	3
With at least one live birth irrespective of current family planning status	0	1	4	9	14	15	20	19	20	12
Total number of couples	34	603	639	601	530	637	562	670	522	4,798

(a) Includes sterilization throughout.
Source: Taiwan IUD-III Survey, 1968, Jain and Sun, 1970.

Table 12. 22. *Current family planning status for all couples and those with IUD not in situ, Taiwan, 1968*

Current family planning status	*All couples (percentage)*	*Couples with IUD not in situ (percentage)*
IUD in situ *at interview*		
Original device	40	—
Reinserted	11 (a)	—
IUD not in situ *at interview*		
Currently using other contraception or sterilization	21 (a)	43
Not currently protected by contraception or sterilization:		
but all pregnancies aborted	5	9
never pregnant	13	26
had at least one live birth	7	15
currently pregnant (only one pregnancy)	3	7
Total percentage	100	100
Number	4,789	2,342
Percentage of women who had at least one live birth irrespective of their current family planning status	12	20
Per cent who are:		
Currently protected by an IUD or contraception or sterilization	72	43
Currently protected or aborting all pregnancies	77	52
Currently protected, aborting all pregnancies, or never pregnant	90	78
Not protected in any of these ways and had at least one live birth	7	15
Not protected in any of these ways and currently pregnant (only one pregnancy)	3	7

(a) Includes women who had a live birth after terminating the first IUD.
Source: Jain and Sun, 1970.

concerns. As Freedman and Takeshita (1963: 303-304) pointed out, the fertility of IUD acceptors depends not only on the births prevented by the first IUD inserted, but

> "also depends on the effect of reinserted IUDs and of the various alternative methods adopted by acceptors who gave up an IUD entirely: sterilization,

abortion, and other methods of contraception. It has been found that those who became acceptors in the program are sufficiently interested in controlling family size to manage to keep their fertility relatively low whether or not they retain the initial IUD."

Table 12.21, which was derived from the third Taiwan island-wide IUD follow-up survey data, shows that 48 months or more after the initial IUD insertion, more than 40 per cent of acceptors were still using the original or a reinserted IUD, 71 per cent were currently using some form of contraception (including sterilization) and 80 per cent had been successful in avoiding any live births. Taking contraception (including sterilization) and abortion together, more than 85 per cent of all IUD acceptors have been protected from having a live birth throughout various times of observation after initial IUD acceptance.

Since acceptors were sufficiently motivated to control their fertility by one means or another, the rather high IUD termination rate was compensated for by adoption of other contraceptive methods and abortion. As Table 12.22 shows, 78 per cent of the third IUD follow-up survey (1968) sample who were not using an IUD at the time of the survey were protected from having a live birth. This proportion was over 70 per cent for all the IUD terminators who were interviewed at different time periods after IUD acceptance.

It is worth noting that in the third IUD survey sample, although only 8 per cent of the 24,984 pregnancies occurring to the sample *before* IUD insertions were terminated by induced abortions, 51 per cent of the 1,576 pregnancies occurring *after IUD* acceptance were so terminated.

The net result of this additional protection of IUD acceptors was a significant drop in post-insertion fertility. The fertility of the third IUD survey sample decreased dramatically from 403 per thousand before acceptance (3 years average) to 75 per thousand in the second year after acceptance, i.e., a drop of 81 per cent. As Table 12.23 shows, the fertility decline after IUD insertion varies with age at acceptance, ranging from 65 per cent of those younger than 25 to 98 per cent of those aged 40 and above. The decline, of course, is the total (or gross) effect of an IUD program, including both the effect of the IUD itself and also the effect of other fertility control measures taken after IUD termination. The problem remaining is to determine how much the acceptors would have reduced their fertility even without the program. Considering that other contraceptives are increasingly available and that people's attitude towards family planning has

been changing favorably, the extent of family planning practice without an IUD program would have been increasing considerably. It should be noted, however, that an IUD program also has an impact on attitude change and that more than a third of the married women of child-bearing age had been recruited to the IUD program with the gross effect of reducing their fertility by 81 per cent, quite a substantial amount.

Table 12.23. *Fertility of IUD acceptors before and after acceptance, Taiwan third IUD survey, 1968*

	Age group					
Fertility per 1,000	*16-24*	*25-29*	*30-34*	*35-39*	*40-49*	*Total*
Before acceptance (3 year average)	566	503	364	368	182	403
After acceptance (2nd year)	200	112	41	27	4	75
Difference:						
fertility	366	391	323	341	178	328
% decline	64.7	77.7	88.7	92.7	97.8	81.4

Source: Taiwan IUD-III Survey, 1968.

On the basis of life table estimates made by K. H. Su (following the method used by Freedman *et al.*, 1971), the proportion of acceptors in the third IUD survey sample having a live birth rose in successive periods after insertion from 5.5 per cent at 18 months to 10.2 per cent at 24 months, 17.1 per cent at 36 months, 22.6 per cent at 48 months and 26.7 per cent at 54 months. Considering only the IUD terminators, the proportion with a live birth rises from 9.2 per cent at 18 months to 16 per cent at 24 months, 26.8 per cent at 36 months, 34.7 per cent at 48 months and 41.5 per cent at 54 months. The higher rates for the terminators are to be expected since they are less likely to be protected, and a disproportionately large number of them were using the IUD to postpone their next child rather than to stop family growth. About 84 per cent of the women who accepted the IUD from 1964 to 1969 said that they were doing so to stop any further family growth, i.e., they already had enough or more than enough children. About 16 per cent of the acceptors were using the IUD to space their children, so that it should not be surprising that they would go on to have additional births. As Freedman *et al.* (1971) pointed out,

"It is a little strange that programs interested in getting more couples to use contraception for spacing often count it as a 'failure' when such couples stop contraception and have another child."

An estimate based on the life table method shows that couples who have had all the children they want are remarkably successful in keeping their fertility rates very low: only 16 per cent had a child 48 months after insertion. On the other hand, among those who still wanted more children, the cumulative proportion having a child increased from 1.5 per cent at the end of the first year to 27.8 at the end of the second year, 49.3 at the end of the third year and 72.0 at the end of the fourth year. It should be noted that 28 per cent of those who said they wanted more children did not have one even 48 months after insertion. Probably some of them had changed their minds about having additional children (or had given an incorrect response originally) or were unable to conceive when they wanted to do so. The figures by Freedman *et al.* (1971), quoted in Table 12.24, show that the post-insertion fertility of women who have accepted an IUD depends on whether or not the women want more children and on whether or not they have had the IUD removed.

According to the estimate made by Freedman *et al.* (1971), the observed post-insertion fertility for acceptors who had never used contraception before the insertion (about 75 per cent of the IUD-II

Table 12.24. *Comparison of actual and expected fertility of IUD acceptors by purpose of acceptance and current use status, Taiwan second IUD follow-up survey, 1966*

Current use status and purpose of acceptance	*Actual proportion with live birth in 18 months*	*Expected proportion with live birth in 9 months, based on age-specific marital fertility in Taiwan, 1966 (a)*
All acceptors	0.05	0.18
Want more children	0.17	0.26
Want no more children	0.03	0.16
All terminators	0.11	0.20
Want more children	0.22	0.27
Want no more children	0.07	0.17

(*a*) 9 months of exposure were subtracted to adjust for the fact that women are not pregnant at the time they accept an IUD, even though this is an overly conservative adjustment for this factor.
Source: Freedman *et al.*, 1971.

sample) was 66 to 77 per cent lower than the "expected" fertility estimated on the basis of prior fertility of the acceptors, and is also 62 to 74 per cent lower than the "expected" fertility of the general Taiwan population of married women. The post-insertion fertility decline of 65-80 per cent as compared with the expected values without contraception is in the same range as other comparisons in Taiwan and Korea (Potter *et al.*, 1968) ([11]). The fact that the IUD acceptors in Taiwan were sufficiently motivated to avoid unwanted live births is also well documented in the matching study mentioned in the last section (Chang *et al.*, 1969).

2. *PILL.* According to the results of the first island-wide pill acceptor follow-up survey (1969),

> "Of the 1,464 cases who stopped use of the first segment of pill, 56 per cent either shifted to other methods (35 per cent) or had begun pill use again (21 per cent). Of those who terminated the first segment of pill use (1,464), 24 per cent had at least one pregnancy and 0.7 per cent were not certain about pregnancy. The average length of time since termination of the first segment was 7 months. Excluding 103 cases who were still pregnant at the time of interview and 8 whose pregnancy was uncertain, 68 per cent of the pregnancies occurring after pill termination ended in induced abortion, and only 25 per cent ended in a live birth by the time of interview. It was not known how the 103 other pregnancies would terminate but live birth was most likely for those pregnant for longer periods. This finding is similar to that of the sample of loop follow-up surveys: many women who start contraception and become pregnant tend to take action (abortion) to prevent births."
>
> (Chen *et al.*, 1970.)

ANALYSIS OF FERTILITY TREND

As mentioned in the introduction, Taiwan's fertility had started to decline steadily before the implementation of the island-wide family planning program. This complicates the measurement of program impact using fertility trend data.

Assuming that the pace of changes in factors affecting fertility, such as proportion married, adoption of fertility control measures in the absence of a program, etc., was the same during the periods before and after the expansion of the program ([12]), the effect of the program on fertility should appear as an accelerated fertility decline which can be measured as the difference in expected and actual fertility of the population.

([11]) The data for Korea have been obtained by John Ross.

([12]) This assumption was made for the convenience of analysis. In fact, it is not known whether the effect of modernization on fertility was accelerated after 1963 or not. This note applies to all of the following sections.

Examination of the trend in the crude birth rate between 1958 and 1971 indicates that there is a significant difference in the trend before and after 1963, when the program was expanded to about 120 of the 360 townships in Taiwan (Figure 12.3).

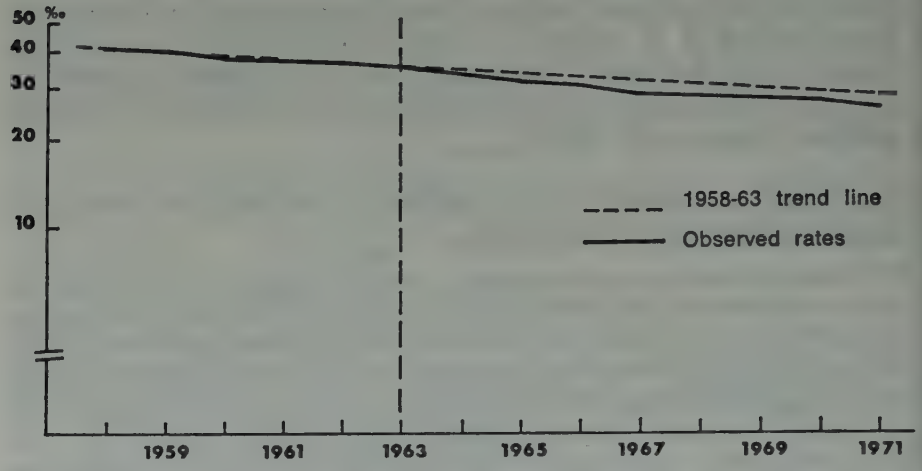


Fig. 12.3. *Crude birth rates, 1958-71, Taiwan area*

The trend lines are estimated using the least squares method and are expressed in the following form:

$$Y_n = Y_0 \ (1-R)^n \qquad (1)$$

Y_0 is the crude birth rate at the beginning of the period under observation taken from the trend line (not actual rate).

Y_n is the crude birth rate n years later.

R is the average annual rate of decline of the crude birth rate.

Then the trend lines of birth rate before and after 1963 can be shown as follows:

For 1958-1963: $Y_n = 41.39 \ (1-0.0302)^n \qquad (2)$

For 1963-1971: $Y_n = 34.39 \ (1-0.0386)^n \qquad (3)$

It is apparent from the formulae that the annual rate of decline of the birth rate for the period 1963-1971 is about 28 per cent greater than that for the period 1958-1963. Therefore, under the assumption made above, it is possible that the family planning program had accelerated the decline of the crude birth rate by 28 per cent annually after 1963.

The difference between the two crude birth rates in 1971 obtained using the two annual rates of decline, i.e., 0.0302 and 0.0386, was 1.83 per thousand points, which is 19.5 per cent of the total decline in the crude birth rate between 1963 and 1971 (9.4 per thousand points on the trend line). The critical question here, however, is the validity of the assumption. As will be discussed in the following section, the components of fertility change may change over time. For example, we know that the age composition of women in the child-bearing period has been changing during the period under observation. This factor alone is sufficient to change the crude birth rate significantly. In order to eliminate the effect of changes in age composition, the total fertility rate can be used in place of the crude birth rate.

As shown in Figure 12.4, the trend lines for total fertility before and after the expansion of the program (1955-1963 and 1963-1971) are quite different.

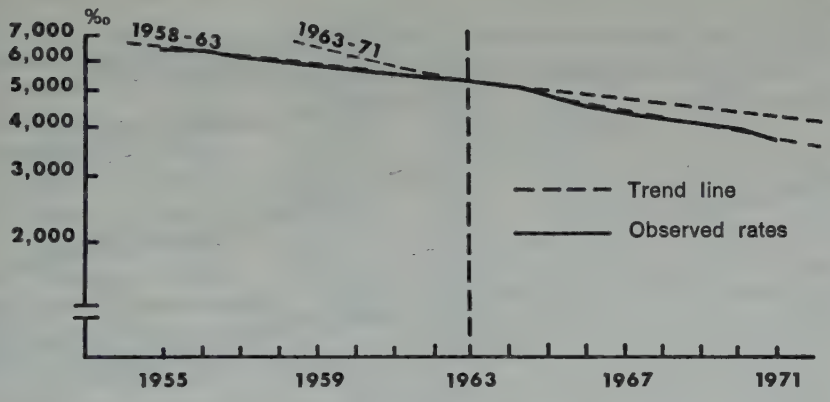


Fig. 12.4. *Total fertility rates, Taiwan area*

Using formula (1) above, the trends of decline in total fertility for these two periods can be shown as follows (replacing crude birth rate by total fertility rate):

For 1955-1963: $Y_n = 6513\ (1-0.0251)^n$ (4)

For 1963-1971: $Y_n = 5269\ (1-0.0426)^n$ (5)

where

Y_n is the total fertility rate n years after 1955 in formula (4), and n years after 1963 in formula (5).

It is clear from these two formulae that the annual rate of decline in total fertility is about 70 per cent (or 1.75 percentage points) greater in the period after 1963 than in the period before 1963. The difference between the two total fertility rates in 1971 obtained using the annual rates of decline, i.e., 0.0251 and 0.0426, was 580 per thousand points, which is 37.4 per cent of the total decline in total fertility between 1963 and 1971 (1,549 per thousand points based on the difference in trend lines). Therefore, if it is assumed that the rate of change in the marriage pattern and other socio-economic factors after 1963 was maintained at the same rate as before 1963, then the decline in total fertility after 1963 was accelerated by the family planning program by 70 per cent per year, and by the end of 1971, about 37 per cent of the decline in total fertility was probably due to the effect of the family planning program.

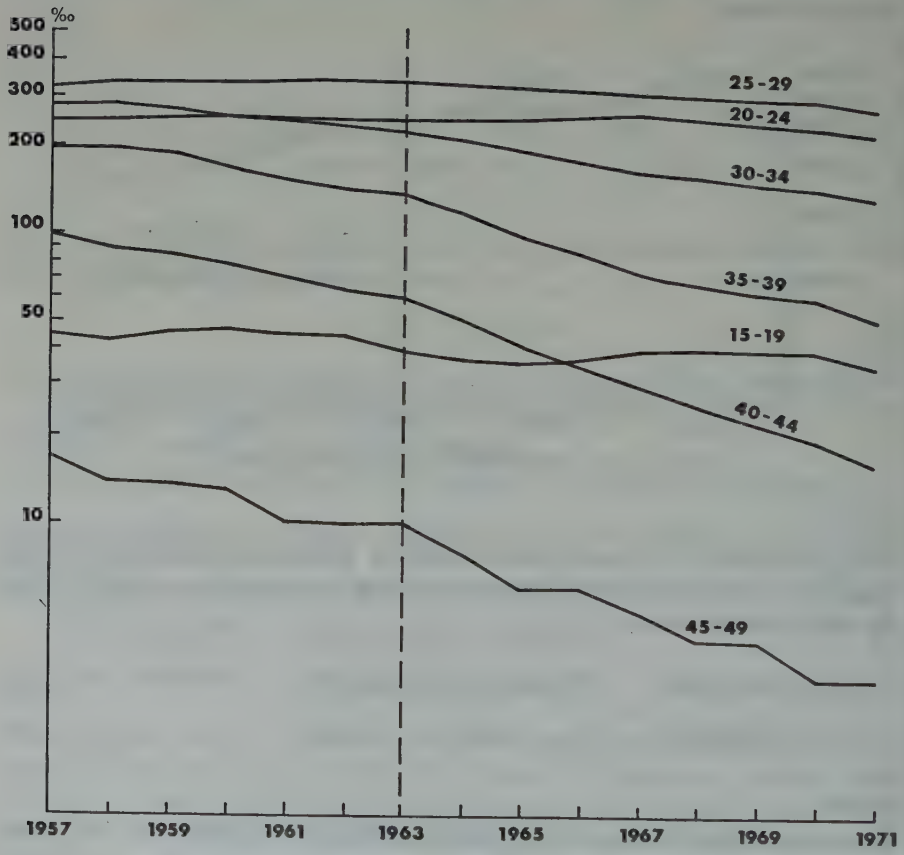


Fig. 12.5. *Age-specific fertility rates, Taiwan*

Table 12.25. *Age-specific fertility trends for all women and for married women the period before (1957-1963) and after (1963-1971) the family planning program, Taiwan area*

e up	1957-1963 (a)			1963-1971			$\frac{(R_1-R_2)}{R_1}$
	Y_{57}	Y_{63}	R_1	Y_{63}	Y_{71}	R_2	×100
				For all women			
9	45.67	43.75	−0.0071	38.89	38.89	0	100
4	250.48	253.24	0.0018	263.76	237.76	−0.0129	—
9	329.87	340.13	0.0051	339.56	281.32	−0.0232	—
4	282.67	229.33	−0.0343	221.73	128.93	−0.0655	−91
9	204.58	136.84	−0.0648	126.44	42.88	−0.1264	−95
4	94.82	60.32	−0.0726	54.98	12.58	−0.1684	−132
9	16.11	9.03	−0.0919	8.71	2.16	−0.1600	−74
				For married women			
9	381.40	362.20	−0.0128	359.73	520.93	0.0474	—
4	403.00	425.00	0.0134	433.79	478.43	0.0123	8
9	374.00	381.60	0.0050	382.11	320.11	−0.0219	—
4	292.60	247.40	−0.0411	239.78	137.78	−0.0669	−63
9	206.80	149.20	−0.0784	139.02	45.42	−0.1305	−66
4	99.40	69.00	−0.0872	63.33	13.33	−0.1770	−103
9	17.40	11.00	−0.1083	10.64	2.91	−0.1496	−38

(a) Trend lines for married women before the program are based on data for 1959 and 1963 due to the of data for 1957 and 1958.

Trend lines for each age group are expressed as follows:

	Before program	*After program*
For all women	$Y_n = Y_{57}(1-R_1)^n$	$Y_n = Y_{63}(1-R_2)^n$
For married women	$Y_n = Y_{59}(1-R_1)^n$	$Y_n = Y_{63}(1-R_2)^n$

Where Y_n is the fertility n years after the beginning of the period, i.e., 1957, 1959, or 1963. For example, ecific fertility for married women in 1961 is expressed as:

$$Y_{1961} = Y_{59}(1 - R_1)^2$$

where "R_1" is annual rate of decrease for each age group.

Source: Taiwan Demographic Fact Book, Ministry of the Interior, 1964 and 1973.

The same analysis can be carried out for each of the age-specific fertility rates to identify the age groups which were affected by the program. The rates are shown in Figure 12.5 and the analysis in the upper panel of Table 12.25. If the assumption that "the rate of change in the marriage pattern and other socio-economic factors after 1963 remained the same as that before 1963" holds, then it may be said

from these results that the family planning program has had a great impact on almost all age groups except 15-19. For age groups 20-24 and 25-29, the rate of change turned from positive to negative. For age groups 30-34, 35-39, 40-44, and 45-49, the rate of decline almost or more than doubled. We know, however, that the marriage pattern in Taiwan has been changing over this period. Therefore, these results must be modified.

In order to eliminate the effect of marriage on fertility, the same analysis can be carried out for the age-specific fertility of *married women.* The results are shown in Figure 12.6 and the lower panel of Table 12.25. The patterns observed here are quite similar to those observed for all women except that the degree of change is more moderate. The fertility of married women in age group 15-19 increased after 1963 and the fertility decline for married women aged 20-24 was very small; the change in fertility of women aged 25-29 turned from positive to negative and the rates of fertility decline in other age groups all accelerated significantly (ranging from 38 to 103 per cent). Since these trend lines are free from the influence of two important demographic components which affect the general fertility level (age composition and the marriage pattern), they show the influence of fertility control measures adopted by married couples. Therefore, the facts are consistent with, but not demonstrative of, a program effect—i.e., the fertility decline was accelerated considerably after the expansion of the program, probably due to additional fertility control measures provided by the program, assuming that the degree of the adoption of fertility control measures after 1963 in the absence of the program continued to rise at the same rate as before 1963. Again, this assumption may not be valid because the adoption of fertility control measures might have been accelerated by the increased degree of modernization, an increased flow of information, a heightened interest on the part of the population and availability of contraceptives in the absence of the program (Mauldin, 1968). However, there is no way of measuring this hypothetical trend, although it is important to note that 71 per cent of IUD acceptors used the IUD as a first method. On the other hand, as Mauldin (1968) has pointed out,

> "the indirect credits (of modernization) may not all be positive, however. If fertility were to rise among some segments of the population—because of better health, better nutrition, etc.—the program effects might well be larger than the difference between actual and expected number of births, based on past trends."

This point seems to be exemplified in the fertility of married women

Table 12.26. *Comparison of the number of births averted by the program estimated from use of program methods and from differences in expected and actual fertility, 1965-1971, Taiwan area*

ear	Number of births averted by the program		Difference	
	Estimated from use of contraceptive methods provided by the program	Estimated from expected and actual marital fertility (a)	(2)−(1) Number	(2)−(1) / (2) %
	(1)	(2)		
	4,841	21,406	16,565	77.4
	18,051	32,191	14,140	43.9
	34,273	43,483	921	2.1
	51,861	41,313	−10,548	−25.5
	68,634	47,330	−21,304	−45.0
	82,986	48,265	−34,721	−71.9
	101,113	59,732	−41,381	−69.3
	361,759	293,720	−68,039	−23.2

(a) The estimation was made for each age group separately based on the expected fertility which was derived the fertility trend before 1963 (1959-1963) as shown in the lower panel of Table 12.25 and the actual ty of the year. The age groups 15-19 and 20-24 were excluded from the calculation because the marital fertility se two age groups had an upward trend during the period 1963-1971.

Source: Table 12.19.

younger than 25, which has either risen significantly or remained unchanged over the past decade even though a significant proportion (5 per cent for age 15-19 and 18 per cent for age 20-24) (see Table 12.3) of women in these age groups have accepted the loop to control their fertility.

One way of testing the assumption that "the degree of adoption of fertility control measures after 1963 in the absence of the program would continue to rise at the same rate as before 1963" is to calculate the number of births averted by the program based on the expected (by extending the fertility trend before 1963) and actual age-specific marital fertility for married women (the lower panel of Table 12.25), and comparing the results with those shown in Table 12.18, which were estimated from the number of current users of program methods. This comparison is made in Table 12.26. It seems that the calculation based on the difference in expected and actual age-specific marital fertility overestimates the number of births averted for the period before 1967 and underestimates the number after 1967.

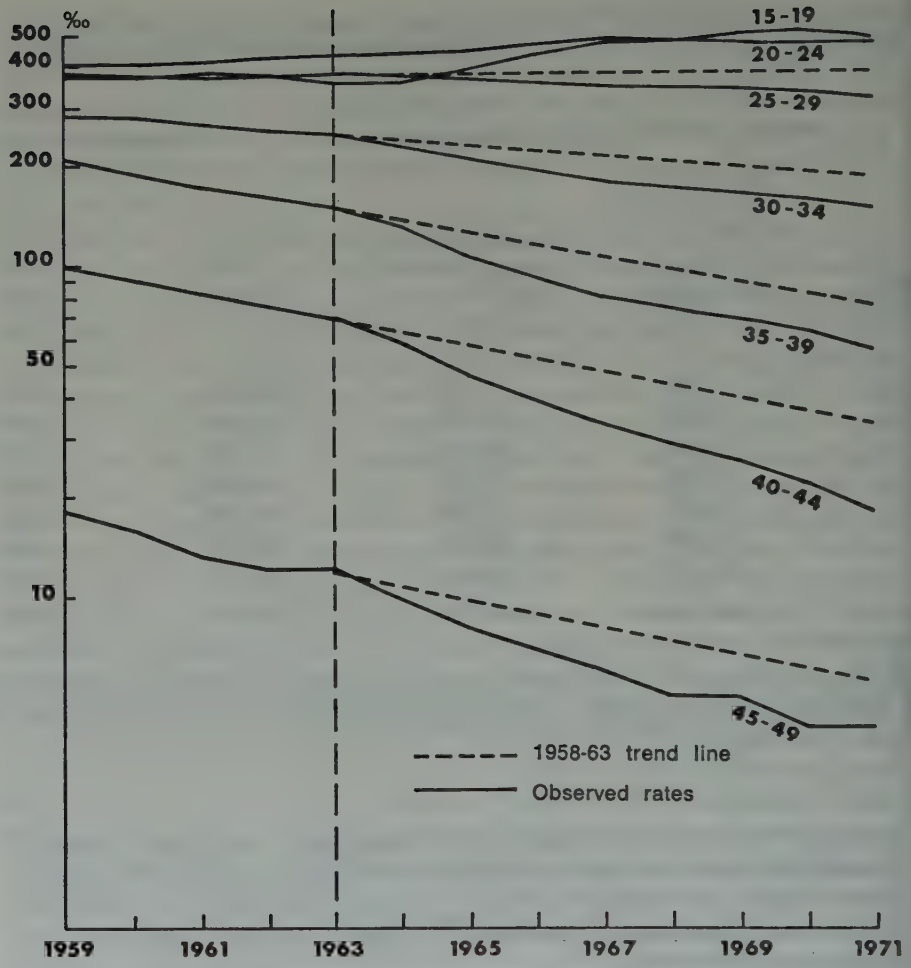


Fig. 12.6. *Age-specific-marital fertility rate, Taiwan area*

Or, to put it another way, the calculation based on the current use of program contraceptive methods underestimates the number of births averted before 1967 and overestimates the number after 1967. However, judging from all factors concerned, it seems that the former situation is more likely, i.e., calculation made on the basis of use of program contraceptive methods is more accurate. If this is the case, then the assumption made above that "the degree of adoption of fertility control measures after 1963 in the absence of the program would continue to rise at the same rate as before 1963" should be corrected.

The comparison of the difference between the two estimates suggests that the degree of adoption of fertility control measures in the absence of the program seems to differ after 1963; it appears to have been very high right after 1963 but to have become lower as time passed, i.e., the *expected fertility* trend in the absence of the program after 1963 should be curvilinear, with a higher rate of decline at the beginning and a lower rate at the end, so that the difference between expected and actual fertility would be smaller at the beginning and larger at the end. This speculation suggests that there is room to reconsider the assumption that the rate of increase in contraceptive use will continue forever at a constant rate. It might also suggest, however, that a part of the program effect is a substitution for what might have happened in its absence.

COMPONENTS OF FERTILITY DECLINE

As stated in the preceding section, the annual rate of decline in the crude birth rate in the period following the expansion of the family planning program (1963-1971) was 28 per cent greater than that in the period before (1958-1963). We know, however, that the level of the crude birth rate is affected by three factors: age composition, marriage pattern (or proportion married in each age group), and marital fertility ([13]). A family planning program is supposed to have an effect on marital fertility by increasing the use of contraception or by indirectly leading to the use of other fertility control measures. It may have some indirect effects on the marriage pattern but has no effect on the age composition of women of childbearing age over a short period. Therefore, it may be argued that the faster decline of the crude birth rate after 1963 is mainly due to changes in age composition and the proportions married in each age group rather than to an accelerated reduction in marital fertility.

In order to reexamine this question, the reduction in the crude birth rate can be decomposed using standardization techniques to show the effect of these three components on the decline. Due to the limitation of the data ([14]), the calculation was made for three periods: 1959-1963 (before the large-scale program); 1963-1967 (early program period); and 1967-1971 (recent period), each covering four years. The results are shown in Table 12.27.

For the period before the program (1959-1963), 58 per cent of the fall in the birth rate can be attributed to a genuine decline in marital

([13]) Assuming that births outside of marriage are insignificant in numbers, as in Taiwan and most Asian countries.

([14]) Data on age-specific marital fertility are available only since 1959.

Table 12.27. *Decomposition of changes in crude birth rates before and after the program expansion and comparison of per cent changes in crude and standardized rates, Taiwan area*

Crude birth rates and components of change	Before program	After program	
	1959-1963	1963-1967	1967-1971
Crude birth rate, beginning of period	40.49	35.64	29.08
Crude birth rate, end of period	35.64	29.08	25.64
Change in crude birth rate due to:			
Age	1.41	1.18	—1.22
Marriage	0.98	0.71	1.56
Marital fertility	2.83	4.35	2.69
All factors (a)	4.85	6.56	3.44
Per cent decline due to:			
Age	29.10	18.00	—35.50
Marriage	20.20	10.80	45.30
Marital fertility	58.40	66.30	78.20
All factors	107.70	95.10	88.00
Interaction	—7.70	4.90	12.00
Crude birth rate, standardized for age and marriage, beginning of period	37.66	31.29	26.39
Average annual change (b)			
Crude birth rate	1.21	1.64	0.86
Standardized birth rate	0.71	1.09	0.67
Per cent decrease per year (c)			
Crude birth rate	3.00	4.60	3.00
Standardized birth rate	1.80	3.10	2.30
Per cent decrease by component			
Age	3.48	3.31	—4.20
Marriage	2.42	1.99	5.36
Marital fertility	6.99	12.21	9.25
All factors	11.98	18.41	11.83

(a) Detail will not add to total because a small part of the total change is due to the interaction of two or more factors.
(b) Obtained by dividing total change by number of years in period covered.
(c) Obtained by dividing average annual change by crude birth rate at beginning of period.

Source: Taiwan Provincial Government. (Following the method used in Freedman *et al.*, 1972a).

fertility. The contribution of the decline in marital fertility to the fall in the birth rate increased to 66 per cent in the period 1963-1967, and to 78 per cent in the period 1967-1971. This indicates that the decline of the crude birth rate after the expansion of the family planning program was not only the result of changes in age composition or proportions married but was also due to the accelerated decline in marital fertility.

As the last panel of Table 12.27 shows, the proportion of the decrease in the crude birth rate due to a decline in marital fertility was 6.99 per cent in the period 1959-1963, increased to 12.20 per cent in the period 1963-1967, but dropped to 9.25 per cent in the period 1967-1971. The annual percentage decline in the birth rate (standardized for age and marriage) is also much higher in the period after 1963, the percentages for the periods 1963-1967 and 1970-1971 being 72 and 28 per cent respectively higher than that for the period 1959-1963. These percentages confirm that the fall in the birth rate after 1963 was largely due to an accelerated decline in marital fertility. Of course, as mentioned earlier, the decline in marital fertility is not entirely due to the effect of the program. As long, however, as the speed of decline in marital fertility after 1963 exceeds that before 1963, it is possible that the family planning program has had an impact on fertility, assuming that the speed of modernization was about the same before and after 1963.

However, as Freedman *et al.* (1972a) point out,

> "these standardization procedures may be misleading. In standardizing to 1959 proportions married, it was assumed that the young unmarried women in 1959 would have had as high fertility, had they been married. The effects of marriage change are exaggerated in so far as those who postpone marriage are also likely to be those with lower fertility."

V. OTHER PROGRAM ANALYSIS

AREAL ANALYSIS OF THE EFFECT OF ACCEPTANCES ON FERTILITY

One attempt has been made by Hermalin (1971) to measure the impact of a program on the fertility of local areas, isolating the effect of some modernization factors, such as level of education, occupation and child mortality. The question raised was whether the family planning program had produced a significant effect on fertility decline above and beyond that attributable to the on-going socio-economic and demographic processes.

The analysis utilized data on the 331 administrative subdivisions ([15]) (townships) of Taiwan in an areal multivariate regression analysis using a model which postulated that an area's level of socio-economic development, its demographic position in terms of age at marriage and child mortality, and its level of IUD acceptances from the family planning program are possible influences on its fertility level. The analysis was carried out using total fertility and age-specific marital fertility in four key age groups as dependent variables for each of the years from 1966 to 1969.

The effects of cumulative IUD acceptance rates on age-specific marital fertility, after taking into account the effects of children's education, proportion married, proportion of males in agriculture and fishing and child mortality are shown in Table 12.28. The results of the analysis are summarized by Hermalin as follows:

> "Over all the 331 areas, a significant negative relation between acceptance rates and *each* fertility measure was found in *each* of the years, with the exception of two years at ages 20-24. In addition, when the regressions were performed among three subsets of areas classified by their level of pre-program fertility (1963), the magnitude of the relation was substantially strengthened in at least one of the subsets. Moreover, the strongest relations were found in those subsets where they might be expected to occur on the basis of the characteristics of the areas in the subset."
>
> "The results by subset strengthen the likelihood that acceptance of IUD's is not merely a substitution for other forms of contraceptive use. The reasoning is as follows: Division of the areas by fertility prior to the program creates areas that should be roughly homogenous in their level of pre-program contraceptive use. If acceptances from the program reflected only substitution for existing contraception, or differences among areas in the mix of program and private contraceptive acceptance, one would not expect differences in acceptance rates to be associated with differences in subsequent fertility (assuming program and private contraception are equally effective). The analysis did *seem* to identify classes of areas and age groups in which program acceptance was serving as a substitute for other contraception practice (low fertility areas). But it also *seems* to identify classes in which program acceptance was serving to decrease fertility through family limitation and spacing (high fertility areas and ages 25-29)."

Even though there are limitations in the data analysis method and some questions regarding the model, as were discussed by Hermalin in his analysis report, it seems that the family planning program did have an effect on fertility above and beyond that attributable to the on-going modernization trend.

The effect of a family planning program on fertility cannot be represented only by acceptance rates because a program has more

([15]) 30 of the 361 townships, which are aboriginal areas, were excluded from the analysis.

Table 12.28. *Beta coefficients for age-specific marital fertility and total fertility from cumulative IUD acceptance rates (per 1,000 married women) in 331 local areas (divided into three fertility groups), Taiwan area, 1966-1969*

ear	Fertility in 1963 (a)	Age-specific marital fertility				Total fertility
		20-24	25-29	30-34	35-39	
	Low	0.05	−0.21 (b)	−0.12	0.04	0.06
	Medium	−0.15	−0.07	−0.14	−0.18	−0.05
3	High	−0.20	−0.15	−0.49 (b)	−0.28 (b)	−0.40 (b)
	Total	−0.09	−0.11 (b)	−0.21 (b)	−0.18 (b)	−0.19 (b)
	Low	0.15	−0.31 (b)	0.05	0.07	0.06
	Medium	−0.22 (b)	0.01	−0.15	−0.12	0.07
7	High	−0.43 (b)	−0.22 (b)	−0.43 (b)	−0.29	−0.33 (b)
	Total	−0.14 (b)	−0.15 (b)	−0.20 (b)	−0.17 (b)	−0.17 (b)
	Low	0.12	−0.08	0.06	−0.19	0.12
	Medium	−0.22 (b)	−0.11	−0.16	−0.09	−0.06
3	High	−0.35 (b)	−0.42 (b)	−0.49 (b)	−0.33 (b)	−0.36 (b)
	Total	−0.10	−0.13 (b)	−0.21 (b)	−0.21 (b)	−0.17 (b)
	Low	0.00	−0.21 (b)	0.08	−0.08	−0.09
	Medium	−0.26 (b)	−0.11	−0.06	−0.14	−0.12
9	High	−0.34 (b)	−0.19 (b)	−0.32 (b)	−0.23 (b)	−0.24 (b)
	Total	−0.17 (b)	−0.14 (b)	−0.16 (b)	−0.14 (b)	−0.15 (b)

(a) Definitions of subsamples:
For age-specific fertility: 331 local areas (townships) were divided into three roughly equal divisions according marital fertility rate of that age group in 1963. For the 35-39 age group, the division was according to l fertility for age group 30-34 in 1963.
For total fertility: 331 local areas were divided according to total fertility in 1963:
≤ 528 per thousand (N = 107);
m = 529-605 per thousand (N = 129);
= > 606 per thousand (N = 95).

(b) Beta coefficients (the partial regression coefficients in standard form) which are two or more times standard error. The cumulative acceptance rate is the cumulative acceptances of IUDs from the family planning m up to March of the year in question or the preceding December, per 1,000 married women. Other independent les which have been included in the multiple regressions are: (1) percentage of females aged 12 to 19 re currently attending school (1965); (2) percentage of male labor force (12 and over) engaged in agriculture shing (1963); (3) average percentage of women currently married at age 20-24 in 1961 and 1963; (4) weighted e of death rates for ages 0-4, in 1961 and 1963.

Source: Hermalin, 1971.

components, such as its educational functions, which will influence married couples to take up fertility control measures not provided in the program, such as sterilization or even induced abortion, and therefore, lead to further reduction in marital fertility. Hermalin has already

Table 12.29. *Percentage who had ever used contraception by selected modernization measu for married women aged 22-39, 1965, 1967 and 1970, Taiwan area*

Modernization measures (a)	Percentage who had ever used contraception (b)			Number of cases		
	1965	1967	1970	1965	1967	1970
Education						
None	19.4	32.9	50.9	1,449	1,525	839
Primary	31.9	42.3	53.7	1,311	2,135	1,325
Junior high	51.4	58.3	71.4	181	289	194
Senior high or more	60.2	65.5	79.2	108	196	133
Urbanization						
Rural township	21.4	32.9	48.2	1,227	1,628	1,004
Urban township	25.7	40.1	53.8	902	1,240	753
Small city	31.7	44.0	62.2	303	333	199
Large city	43.3	55.4	69.1	617	944	535
Newspaper reading						
Daily	58.4	61.4	73.5	305	593	356
Frequently	40.9	49.2	61.0	132	333	95
Seldom	34.6	40.5	57.7	434	596	278
Never/can't read	21.9	35.6	52.0	2,178	2,623	1,762
Mass media exposure						
Low		30.0	47.0		1,254	943
Medium		40.9	58.1		1,862	1,126
High		55.0	71.8		1,029	422
Number of modern objects owned						
0-3	19.4	35.4		1,245	1,207	
4-5	27.3	35.6		770	939	
6-7	28.5	39.2		548	812	
8 or more	51.4	52.5		486	1,187	
Total	28.1	41.1	55.5	3,049	4,145	2,491

(a) Data not available on all measures for each time period.
(b) Including sterilization.
Source: Freedman et *al.*, 1972b.

noted in his discussion on the limitation of the analysis that the model used excluded inputs of fieldworkers and doctors who are involved in the program. It would be interesting to see the analysis performed using an alternative model which included these inputs.

KAP MEASURES OF INCREASE IN CONTRACEPTIVE PRACTICE

An indirect measure of program effect on fertility is its ability to increase contraceptive use among married couples, especially those in less modernized sectors. Taiwan's family planning program has been concentrating its effort mainly on women who are in lower status and high fertility strata.

According to the results of the three KAP surveys (1965, 1967 and 1970), the very large increase in contraceptive use has occurred at every social and modernization level, but the increase has been especially great among the less modern: the less educated, those in rural areas, those with low mass media exposure, those who do not read newspapers, and those who are least modern consumption oriented (Freedman *et al.,* 1972b). Therefore, the social differentials in the use of contraception have narrowed considerably (Table 12.29). For example, because contraceptive use among those with no education more than doubled while the increase among the most educated was about a third, the spread between lowest and highest level changed from 19 to 60 per cent in 1965 to 51 to 79 per cent in 1970. The gap has been narrowed both proportionately and in absolute numbers.

As was also pointed out by Freedman, Coombs and Chang (1972b),

> "with the trend toward more education and urbanization in Taiwan, it would be possible for the increased use of contraception to be merely a reflection of population shifts into the strata that tend to use contraception more, but this has not been the situation."

The proportion who had ever used contraception in 1970, standardized for levels of education in 1965, shows that only a slight part of the increase in use of contraception between 1965 and 1970 was attributable to a rise in educational level. Similar results were obtained when the rates were standardized for changes in urbanization, for marriage duration, and for the number of children ever born. This indicates that the use of contraception in Taiwan is being increasingly accepted by couples who have not yet moved into the more modern strata. We may interpret this as one of the achievements of the program, which has had a significant effect on lowering the fertility of less modernized people.

SOME RELEVANT OPERATIONAL RESEARCH

In order to maximize the number of program acceptances, Taiwan has conducted about a hundred research projects to find out the best approaches and contraceptive methods and to evaluate specific components of the program. Five of these studies are described below ([16]).

1. *THE TAICHUNG STUDY.* Before Taiwan expanded its family planning program on an island-wide scale, many questions were raised regarding the effective operation of such a program. In order to answer these questions, the Taichung Study was designed by Bernard Berelson and Ronald Freedman and was carried out jointly by the Provincial Maternal and Child Health Institute and the Taiwan Population Studies Center from February to October 1963, with the assistance of the Population Council and the Population Studies Center of the University of Michigan. This study has been well documented in a book by Freedman and Takeshita (1969).

The objective of the study was to answer the following major questions about the planning and operation of a family planning program:

(1) To what extent can the practice of family planning be increased by a massive information and service campaign of short duration?
(2) Is it necessary to approach both husbands and wives in an educational program, or is it sufficient to approach the wife alone?
(3) Can family planning ideas be spread cheaply and simply by written communication through the mail?
(4) Can direct communication to systematically spaced subgroups of a population indirectly affect a much larger population by diffusion from the initial foci?
(5) Does a new method of contraception, the intrauterine contraceptive device (IUD), have distinctive advantages in terms of acceptance and diffusion?
(6) If there is significant adoption of family planning, will it accelerate the fertility decline already begun in Taichung and Taiwan?
(7) Which demographic and social characteristics are most important in determining whether a couple will accept family planning in an organized program?

([16]) Most of the studies have been summarized in a recent publication of the Taiwan Provincial Committee on Family Planning (1973).

(8) What are the characteristics of the large proportion of couples who express an intention of accepting family planning but fail to do so?
(9) Was the recent fertility of those accepting family planning high enough that their use of effective contraception would produce a significantly large reduction in birth rates?
(10) Which characteristics of the couples are related to persistence in the effective use of family planning once it is accepted?
(11) How did the discussion and perception of what others were doing about family planning affect information and acceptance?

The essential design of the study included the division of Taichung City, the Taiwan Provincial Capital (population 300,000), into three comparable parts, each part receiving a different intensity of four "treatments", including:

(1) posters and some meetings only;
(2) mailing plus posters and meetings;
(3) all the major stimuli except a personal visit to the husband;
(4) all the stimuli of the program, including visits to the husbands.

Ten clinics were opened in the city to provide contraceptive services on a "cafeteria" basis. Pre- and post-campaign surveys were carried out to evaluate the program. The first effects of the program were observed through March 31, 1964 and the major findings were as follows (Freedman and Takeshita, 1969):

(1) A large information and service campaign of short duration in a developing country can significantly increase the practice of family planning in the population.
(2) It may not be necessary to approach both husbands and wives in a family planning program.
(3) Letters were not effective in increasing the acceptance rate, despite the fairly high literacy rate.
(4) Diffusion played a major part in that the campaign had an effect far beyond the couples directly influenced by the program.
(5) Effective small group meetings had an important role in increasing the acceptance of the IUD.
(6) The new IUD was chosen by a large majority of acceptors.
(7) Taichung's fertility decline accelerated in the year following the experiment, and exceeded that of the other cities or of the province by a considerable margin.

(8) Many families strongly interested in family planning were able to adopt more satisfactory and effective methods.
(9) Such a large-scale effect could be carried out according to plan, with measured results, without political repercussions and in such a way as to provide a secure basis for the much larger island-wide effort that immediately followed it.

The approaches and contraceptive method (the IUD) which worked so well in the study were adopted in the island-wide program which was started in 1964.

2. *FIELD WORKER PERFORMANCE STUDY.* The purpose of this study was to determine from regular program statistics on field workers and from demographic characteristics of local areas (townships) what type of worker, what mix of working procedures, and what factors specific to the area produced the greatest number of IUD acceptances (Chen and Hermalin, 1970).

Factors considered included:

(1) age, education, marital status and length of experience of the worker,
(2) size, fertility level and educational level of the women, proportion of labor force engaged in agriculture and existence of contracted doctors and/or other workers serving the townships,
(3) methods of field work such as the numbers of home visits, group meetings held and coupons issued, and control office assessment of performance.

Conclusions from the 1966 data analysis showed that married workers with experience of the local area, with a moderate education (junior high school) and who issued many coupons, usually during home visits, were the most effective.

3. *STUDY ON FREE LOOP OFFER FOR A LIMITED TIME.* The purpose of the study was to see whether a free offer for a limited time only would effectively promote acceptance of the loop, which is usually offered at NT $30 (US $.75) (Gillespie *et al.,* 1970).

This study was conducted in two rural townships in Changhua County (with a combined population of 50,000) during August-December 1964. The Lippes loop was provided at the two local health stations on two afternoons a week for a period of three months. In addition, the loop could be obtained from eight private doctors located in a city about 25 kilometers away, and conventional contraceptives could be obtained from local midwives, hairdressers, barbers, village

chiefs, farmers' associations, drugstores, private doctors and health stations in each of the two townships. Six field workers distributed flyers and free coupons to 93 per cent of the 8,000 households in the townships over a six week period

During the first three weeks, 1,140 couples (20 per cent of the target population) accepted family planning services. The acceptance rate was higher in these townships than in any other of the 359 townships which provided family planning services. Of these acceptors, 60 per cent selected the loop and almost all of them had it inserted at the local health station. The follow-up study conducted ten months later showed that the continuation rate was about 90 per cent. In this study the referral cost of field worker time was only US $.80 per case compared with about US $2.50 for the island-wide program. Thus the extra payment to the doctor, i.e., the US $.75 normally paid by the acceptor, was more than balanced by the savings in referral costs, which had been lowered by increasing the number of acceptors in a short period. As a result of this study, the free offer of the loop for a limited period was expanded and then became a regular part of the program.

4. *STUDY OF TYPES OF PILLS PREFERRED.* Due to supply and other reasons, in 1967 and 1968 the program had used mostly two brands of pills, Ovulen (containing 1 mg of Ethynodiol diacetate and 0.1 mg of mestranol) and Primovlar (containing 0.5 mg Norgestrel and 0.5 mg Ethynyl estrodiol). Both brands of pills also had two regimens, 21 or 28 pills per cycle. The question arose as to which of the four variations was best (Sun and Lee, 1970).

To answer these questions, an experimental study was carried out in 1968-1969 in three matching groups of townships in Taiwan. Each group of townships (with 100,000-150,000 married women aged 20-44) was offered only one type of pills, Ovulen-21, Ovulen-Fe-28 or Primovlar-21. The acceptors were followed up four months later to compare the use-effectiveness of each type of pill.

The comparisons of continuation rates and side-effects indicated that the lower-dose estrogen pills (Primovlar) were better than those with the higher dose and that 28 pills per cycle were better than 21.

5. *MASS MEDIA RESEARCH.* The two main functions of a family planning program are:

(1) to create a demand for the practice of family planning, and
(2) to provide services promptly to those who need them.

Before the announcement of a family planning policy in 1968, the first function which is primarily educational was carried out mainly in the form of face-to-face contact (home visiting by field workers) and group meetings. Since then, however, mass media programs have started to play an important role; their share of the total family planning budget has increased from 6 per cent in 1968 to 10 per cent in 1969 and 22 per cent in 1970, then dropped to 14 per cent in 1971, 7 per cent in 1972 and 11 per cent in 1973. Mass media activities include radio and television programs, newspapers and magazine advertisements and articles and the distribution of printed materials. The activities are especially strong in November each year which since 1971 has been designated Family Planning Month. The educational content has also changed, from simple information on contraceptive methods and their provision to information on the advantages of the small family and of child-spacing.

Research into the choice of the best media and their effects began to interest the program administrators in 1966. The first of the large-scale research programs was the study carried out in Kaohsiung City, Taiwan's second largest city (population then 660,000) in the period 1966-1968 to study the increase in loop acceptance through the more active use of mass media, extensive use of an organizational network and increased work input. The study has been described in detail by Cernada and Lu (1972). One of its important findings, that public information channels, particularly radio, can get family planning information to married women at a comparatively low cost, was instrumental in getting the island-wide mass media campaign started in 1968.

Another important research project was the evaluation of the 1971 Family Planning Month mass media activities through an interview survey of 2,013 women in the age group 18-34. The results indicated that the activities during the month had been able to get information to about 44-65 per cent of the target population. Television had been the most important channel in reaching the women, followed by posters, radio and newspapers in that order. It also made clear that use of a multi-media approach is essential if most women are to be reached. Evaluation of the effectiveness of Family Planning Month in terms of an increase in acceptance was difficult because in November and December 1971 free loop insertions, which always produce an increase in acceptors, were offered island-wide. There were significant increases in acceptances in November and December for all three methods, but this increase was not maintained past January.

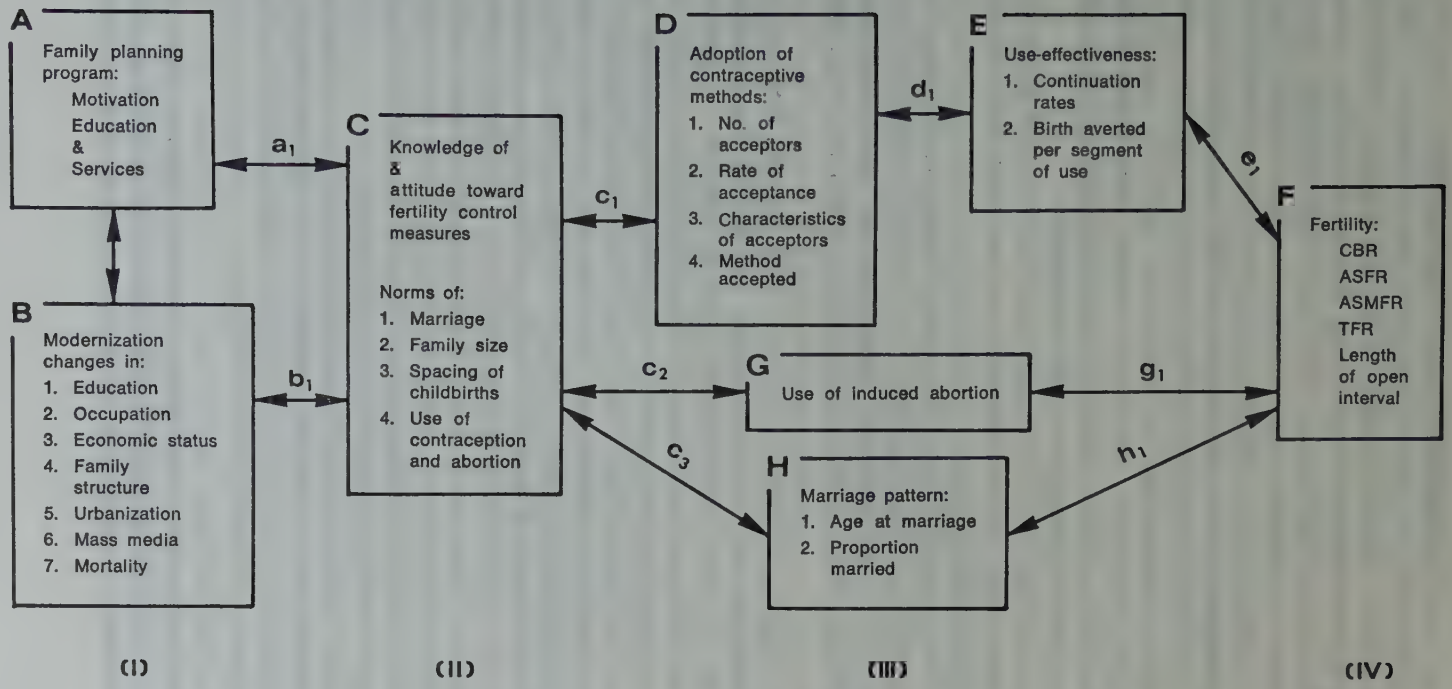


Fig. 12.7. *Factors related to fertility change*

VI. SUMMARY AND DISCUSSION

The goal of Taiwan's family planning program was to reduce the natural increase rate from 30 per thousand in 1963 to less than 20 per thousand in ten years. This goal was translated into a specific number of couples to be recruited for the practice of program methods, mainly the loop, which was supplemented by the pill in 1967 and by condoms in 1970. By the end of 1971, more than 45 per cent of the 1.8 million eligible couples had been recruited for use of these contraceptive methods and by the end of 1972 the rate of natural increase had declined to 19.4 per thousand. How much of this dramatic decline in fertility, however, was attributable to the effect of the family planning program is still in question.

A series of systematic analyses have been carried out to evaluate the program using service statistics, survey data and the excellent household registration data.

For the convenience of discussion, some major factors related to fertility change may be grouped into several categories as shown in Figure 12.7. This scheme follows basically the model presented by Freedman (1963), with specific factors listed for discussion purposes.

The original sources of influence are divided into two groups for convenience of analysis: effects of the family planning program (Box A), and modernization factors (Box B). Other factors are placed between these two groups of source factors and fertility. Box C includes factors related to knowledge and attitude towards fertility control methods and services, and norms about family size, marriage, spacing of births, use of contraception and abortion, and other intermediate variables. Factors in Box C affect the fertility level through Boxes D and E, Box G, or Box H, i.e., adoption and effective use of contraceptive methods, use of induced abortion, or marital status. Age composition which can also affect fertility was not shown in the scheme. Fertility (Box F) could be expressed in the forms of crude birth rate, age-specific fertility, age-specific marital fertility, total fertility, or length of birth intervals.

Most of the influences of these factors are in two directions: forward (toward fertility) and backward (toward source factors). For the purpose of the study, however, only influences of a forward direction are discussed unless otherwise noted.

Even though both the family planning program (Box A) and modernization factors (Box B) can affect fertility through all of the factors between them (Boxes C-H), the question here is how strong

is each channel of influence. To some demographers (e.g. Freedman, 1963; Freedman and Takeshita, 1969), important channels of influence are:

(1) $b_1 - c_1 - d_1 - e_1$,
(2) $b_1 - c_2 - g_1$,
(3) $b_1 - c_3 - h_1$.

For family planning administrators, however, one important channel of influence is:

(4) $a_1 - c_1 - d_1 - e_1$,

and some even think that the following channels exist:

(5) $a_1 - c_2 - g_1$,
(6) $a_1 - c_3 - h_1$.

In order to identify the main channels of influence, we have discussed in this chapter the influence of factors in several boxes.

In section II, the number and rate of acceptance and the characteristics of acceptors were discussed for each method promoted by the program. Since the discussion was limited to three program methods (IUD, pill and condom), the achievements were assumed to have resulted mainly from program effort. It is certain, however, that modernization also had some effect on the adoption of program methods. The data from service statistics show that about 713,000 married women (39.6 per cent of the 1.8 million married women aged 20-44) had tried an IUD (mainly Lippes loop) at least once, 229,000 women (13 per cent) had accepted pills, and 110,000 couples (6 per cent) had accepted condoms. With these three methods taken together, but eliminating the overlapping cases, it was estimated that, by the end of 1971, the program had recruited at least 45 per cent of the 1.8 million married women aged 20-44. Using KAP survey results, one should be able to identify the relative importance of this program acceptance compared with general acceptances. These acceptors were mostly aged around 30, had 3-4 children, had lower education than non-acceptors and came into the program 2-3 years after a childbirth; a large proportion of them had had no previous contraceptive experience. It was concluded that this level of acceptance of program methods should have a significant effect on fertility.

This was followed by a discussion on use-effectiveness of the program methods (Box E). It has been noted that the continuation rate of use was much better for the IUD than for the pill or condom. For the IUD, when longer-period observation data are made available,

it becomes clear that two years after initial insertions, reinsertions tend to balance out the discontinuations so that the proportion of IUD users will be maintained at about 50 per cent of the initial acceptance. This provides a basis for estimating the number of women who are currently using the IUD at a specific time.

Applying Potter's estimation of the number of births averted per first segment of IUD use, it was estimated that about 542,000 births could be averted by 700,000 first IUDs inserted between 1964 and 1971 among married women aged 20-39, an average of .78 births averted per first segment of use. The number of births averted by one year's first IUD insertions ranges between 63,000 and 83,000. It was also noted, however, that Potter's estimation was based on clinical study data which give better continuation rates than are generally true. It should be revised according to the results of the island-wide survey data.

The number of births averted each year by use of the IUD was estimated using the number of current users and the fertility of acceptors. The number of births averted by the pill and condom programs was estimated from the number of pill cycles and condoms distributed. The total number of births averted by the use of these three program methods in each year is shown in Table 12.18. An attempt has been made to differentiate between the number of births averted by the use of program methods and by non-program methods, including induced abortion. The total number of births averted was estimated as the difference between the actual number of births and a theoretical maximum number of births which was estimated by applying a set of age-specific marital fertility rates in the absence of fertility control practice. It was found that the proportion of births prevented by the use of program methods had increased from 11.4 per cent in 1965 to 46.8 per cent in 1972.

The follow-up surveys of IUD acceptors indicate that their fertility continues to be very low even after termination of use of the IUD. This means that once a woman accepts the IUD, she will be sufficiently motivated to adopt other fertility control measures after its removal. This is an effect which goes beyond adoption and use-effectiveness of the program methods and is attributable to family planning as a program effort. It is usually mixed with the effect of modernization and the relative importance of the two factors is difficult to distinguish. The fertility of the third IUD survey sample decreased 81 per cent from 403 per thousand before acceptance (a 3 year average) to 75 per thousand in the second year after acceptance.

An attempt has been made by Chang *et al.* (1969), to measure the difference in fertility of women who were recruited by the pro-

gram and of those who were not but who have characteristics similar o the acceptors. The result shows that, while the fertility of accept- ors declined by 80 per cent after acceptance, the non-acceptors' ertility also dropped 48 per cent. A great part of the difference night be attributed to program influence.

So far, the discussion has been limited mainly to the use of con- raceptive methods provided by the program and to the probable number of births prevented. Very little is known about contraceptive use outside or in the absence of the program even though the point has been touched on in some studies. In other words, the approaches taken so far have tried to measure the effect of the pro- gram directly from its achievement in terms of acceptance and use- effectiveness of the methods accepted (Boxes D and E). We know hat there are more areas in which modernization can influence ertility reduction such as the use of induced abortion or non-program methods, and changes in the marriage pattern.

According to the results of the KAP survey taken in early 1970, 44.3 per cent of currently married women aged 20-44 were using some form of contraception. Of these, 17.0 per cent were using methods provided by the program and 27.3 per cent were using other methods; that is to say that of the total current users, 38 per cent were using program methods. This gives some idea of the import- ance of the program in the total use of contraception. It should be pointed out, however, that many of the couples who were currently using non-program methods had once been program method users and had therefore been motivated by the program.

Freedman *et al.* (1972) have tried to break down the fertility (crude birth rate) decline in order to identify the contribution of the three factors, age composition, proportion married, and marital fer- ility. It was found that an increasingly large proportion of the decline n the birth rate *after* the expansion of the program was due to the decline in marital fertility, which is closely associated with the pro- gram effort. The question still remains, however, as to how much of he decline in marital fertility was due to the program.

An analysis of the various fertility trends (crude birth rate, total ertility rate, age-specific fertility rate, and age-specific marital fer- ility rate, especially the last one), shows that the trend of fertility decline after 1963 was much faster than that before 1963 for most of he groups aged 25 and above. Under the assumption that the effects of modernization on fertility before and after 1963 were constant, this acceleration in fertility decline over the pre-1963 trend could be seen as the degree of additional contribution made by the program.

To sum up, there is considerable evidence that the fertility decline

was to a large degree a result of the reduced fertility of (1) couples practicing program methods and (2) couples who began in the program and then shifted to non-program methods or some combination of the two over time. This evidence indicates at least that the program was an important means by which a massive birth control and demographic change occurred.

It is *possible* that in the absence of the program other means would have been used by this large number of couples to reduce their fertility; there is no way of definitely proving or disproving this point. However, it is also quite possible that without the legitimation and facilitation of the program, the adoption of birth control and the resulting fertility decline would have been much slower. We *know* that couples practiced birth control through the program with considerable effect. Substitution without the program of course would have been possible but the extent to which it would have been feasible will remain unknown.

The likelihood of a significant program effect is supported by the following:

(1) The very rapid adoption of contraception by the general population between 1965 and 1970 occurred at least as rapidly among the rural, less-educated masses as among the more elite part of the population. Quantitatively, it is this population that accounts for most of the fertility decline. It seems unlikely that these rural masses would have had such easy access to methods without the program.
(2) Although it is frequently argued that modernization leads to fertility decline by changing social roles and as a result the desired number of children, this has not been the pattern in Taiwan. For example, between 1965 and 1970 there was little change in desired family size or in the proportion of couples who wanted no more children but there were very large increases among all strata in the proportion practicing contraception among those who wanted no more children.
(3) The areal analyses by Hermalin (1971) and by Schultz (1970) indicate that program inputs do make a difference in fertility after taking into account the social and economic levels of the areas and their prior demographic history.

The studies carried out so far, however, have not been able to identify precisely how much of the fertility decline in the past decade was due to the program and how much to modernization. As pointed out in the text, the available data are not sufficient, and methods of

analysis need revision or improvement. The areas which need more study include Box C and the use of contraception outside the program. Box C will be increasingly important when the changes in attitudes and norms become the key to further fertility decline. When the program moves its effort in this direction, as Taiwan is doing, it would not be appropriate to measure the effect of the program on fertility only through the acceptance of contraceptive methods. The channels of influence, such as: $a_1 - c_2 - g_1$ and $a_1 - c_3 - h_1$ would be strengthened in addition to $c_1 - d_1 - e_1$. These, therefore, require more careful study. Multivariate analyses could be applied more extensively to help clarify the mechanisms of influence of the program or of modernization on fertility.

REFERENCES

Avery, R.

1972a "A brief note on patterns and implications of long-term IUD usage in Taichung, Taiwan," Taiwan Population Studies Working Paper No. 20 (October), Michigan Population Studies Center.

1972b "Calculation of the number of IUDs in place in Taiwan at the end of 1971", memo, March 17.

Berelson, B., and R. Freedman

1964 "A study in fertility control," Scientific American (May): 29-37.

Cernada, G., and L. P. Chow

1969 "The coupon system in an ongoing family planning program," American Journal of Public Health, 59(12): 2199-2209.

Cernada, G. (editor)

1970 Taiwan Family Planning Reader, Taichung, the Chinese Center for International Training in Family Planning (particularly Part V, "The Program Effect on Fertility," pp. 291-351).

Cernada, G., and L. Lu

1972 "The Kaohsiung Study", Studies in Family Planning, 3(8): 198-203.

Chang, M. C., T. H. Liu and L. P. Chow

1969 "Study by matching of the demographic impact of an IUD program: A preliminary report", Milbank Memorial Fund Quarterly, 47(2): 137-157.

Chen, F. L., M. C. Chang, T. H. Sun and G. P. Cernada

1970 "First island-wide pill acceptor follow-up survey, 1968", pp. 225-243, in G. P. Cernada (ed.), Taiwan Family Planning Reader, Taichung; The Chinese Center for International Training in Family Planning.

Chen, H. C., and A. I. Hermalin

1970 "Factors affecting the performance of field workers in the Taiwan IUD program", Taiwan Population Studies Working Paper, No. 7. (Sept.), Michigan Population Studies Center.

Chow, L. P.

1968 "Evaluation of the family planning program in Taiwan, Republic of China", Journal of the Formosan Medical Association, 67: 280-308.

Chow, L. P., R. Freedman, R. G. Potter and A. K. Jain

1968 "IUD termination rates in a mass family planning program: the case of Taiwan", Milbank Memorial Fund Quarterly, 47(2): 215-236.

Council for International Economic Cooperation and Development, Republic of China

1973 Statistical Data Book.

Davis, K.

1967 "Population policy: will current programs succeed", Science, 158(3802): 730-739.

Freedman, R., A. I. Hermalin and T.H. Sun

1972a "Fertility trends in Taiwan: 1961-1970", Population Index, 38(2): 141-166.

Freedman, R., L. C. Coombs and M. C. Chang

1972b "Trends in family size preferences and practice of family planning: Taiwan, 1965-1970", Studies in Family Planning, 3(12): 281-296.

Freedman, R., A. K. Jain, A. I. Hermalin and T. H. Sun

1971 "Fertility after insertion of an intrauterine device in Taiwan's family planning program", Social Biology, 18(1): 48-54.

Freedman, R., and J. Y. Takeshita

1969 Family Planning in Taiwan: An Experiment in Social Change, Princeton; Princeton University Press.

Freedman, R., J. Y. Takeshita and T. H. Sun

1964 "Fertility and family planning in Taiwan: A case study of the demographic transition", The American Journal of Sociology, 70 (July):16-27.

Freedman, R.

1963 The Sociology of Human Fertility, Current Sociology, Vol. X/XI, No. 2, 1961-62.

Gillespie, R., L. P. Chow and H. C. Chen

1970 "Free Loop Offer for a Limited Time," in Cernada (ed.) Taiwan Family Planning Reader, Taichung, The Chinese Center for International Training in Family Planning, pp. 187-191.

Grant, J.

1972 "Taiwan, A Case Study in Development," War on Hunger, Vol. VI, No. 12, U.S. Agency for International Development.

Hauser, P.

1967 "Family planning and population programs," Demography, 4(1): 397-416.

Hermalin, A. I.

1971 "Taiwan: Appraising the effect of a family planning program through an areal analysis", Taiwan Population Studies Working Paper No. 14 (Nov.), Michigan Population Studies Center.

Hermalin, A. I., and L. P. Chow

1971 "Motivational factors in IUD termination: Data from the second Taiwan IUD follow-up survey," Journal of Biosocial Science, 3: 351-375.

Jain, A. K., and T. H. Sun

1972 "Inter-relationship between socio-demographic factors, lactation, and post-partum amenorrhea," Demography India, 1(1): 78-91.

Jain, A. K. and T. H. Sun

1970a "Preliminary analysis report of the third Taiwan IUD follow-up survey in 1968: Termination, reinsertion, current family limitation, and fertility," mimeo., Taiwan Provincial Committee on Family Planning.

Liu, P. T., L. P. Chow and R. Abbey

1972 "A study on IUD retention by curve-fitting," Demography 9(1): 1-11.

Mauldin, W. P.

1968 "Births averted by family planning programs," Studies in Family Planning, 33 (August): 1-7.

Moots, B., and H. C. Chen

1973 "Preliminary Report of Taichung IUD-4 Study", Taiwan Population Studies Working Paper No. 22 (March), Michigan Population Studies Center.

Peng, J. Y., L. P. Chow and L. Corsa

1970 "Medical correlates of termination of use of intrauterine contraceptive device in Taichung," International Journal of Fertility, 15(2):-120-126.

Potter, R. G.

1969 "Estimating births averted in a family planning program," pp. 413-434 in Behrman, Corsa, and Freedman (eds.), Fertility and Family Planning: A World View, University of Michigan Press.

Potter, R. G., L. P. Chow, A. K. Jain and C. H. Lee

1969 "Effectiveness of intrauterine contraception: Termination levels and correlates," pp. 241-273 in R. Freedman and J. Y. Takeshita, Family Planning in Taiwan: An Experiment in Social Change, Princeton, Princeton University Press.

Potter, R. G., R. Freedman and L. P. Chow

1968a "Issues in the evaluation of a family planning program: The case of Taiwan", Industry of Free China, 24(5): 16-28.

1968b "Taiwan's family planning program," Science, Vol. 160, May.

Ross, J. A.

1966 "Cost of family planning programs," pp. 761-762 in B. Berelson *et al.* (eds.), Family Planning and Population, Chicago; University of Chicago Press.

Ross, J. A. and J. E. Forrest

1971 "Program effects upon fertility," to appear in Nathan Keyfitz (ed.), Statistical Problems in Population Research.

Schultz, T. Paul

1970 Evaluation of Population Policies: A Framework for Analysis and Its Application to Taiwan's Family Planning Program, The Rand Corporation, Santa Monica, California.

Sullivan, J. M.

1972 "The influence of demographic and socio-economic factors on infant mortality in Taiwan, 1966-1968," Taiwan Population Studies Working Paper No. 18 (September), Michigan Population Studies Center.

Sun, T. H., and C. H. Lee

1970 "Medical concerns with loop and pill," pp. 245-256 in G. P. Cernada (ed.), Taiwan Family Planning Reader, Taichung, The Chinese Center for International Training in Family Planning.

Taiwan Provincial Department of Cultural Affairs
Taiwan Demographic Fact Book, 1963, 1964, 1971 and 1972.

Taiwan Provincial Committee on Family Planning
1969-1972 Family Planning Reference Book, Vols. I-VIII.
1973 Taiwan Population Studies Summaries, Taichung.

Chapter 13

OVERVIEW

by
Albert I. Hermalin and C. Chandrasekaran

This Overview was prepared with the assistance of Anrudh K. Jain, Dorothy Nortman, Robert G. Potter, K. S. Srikantan, H. Bradley Wells and David Wolfers. Each of the aforementioned supplied a short paper on one or more of the previous chapters. Their comments have been incorporated into this general overview. The editors bear responsibility for the manner in which the ideas of these discussants have been paraphrased or cited in this larger context. Direct quotations from the papers will be cited by author's name.

I. INTRODUCTION

Many of the previous chapters have been devoted to illustrating how a single technique has been used to evaluate some aspect of a family planning program. These chapters provide expositions of the technique, give examples of their application and point out their potential lines of further development as well as their limitations. One objective of the present chapter is to review the rationale and some of the major conceptual and measurement problems associated with each technique. In some instances, additional data or analyses are introduced where these are thought to complement the material in the primary chapter. A third aim is to attempt some synthesis of the various approaches by grouping them in terms of the level of program effect with which they are designed to deal.

Before proceeding directly to this task, it may provide some perspective on the coverage of this volume to elaborate briefly on Chapter 1, which took up some of the broader issues involved in evaluation research of family planning programs. That chapter defined evaluation as the assessment of results achieved by a program directed toward some valued objective. Among the different aspects

of program objectives which may be the focus of an evaluation, Suchman (1967: 61-66) distinguishes five classes:

(1) Effort—measurement of the quality and quantity of activity that takes place.
(2) Performance or Effect—measurement of the results of effort, which can relate to objectives at different levels, e.g. change in knowledge about contraception, adoption of contraception, duration of use, etc.
(3) Adequacy of Performance—measurement of effect in relation to the total need or to the established targets.
(4) Efficiency—evaluation of the cost of the activities in relation to the effect achieved, with consideration of the cost of alternative strategies.
(5) Process ([1])—determination of the means by which a program achieves its results.

The chapters in this volume do not illustrate all the objectives of a family planning program which are subject to research evaluation, nor do they cover all the potentially useful evaluative approaches. Among the aspects of a program subject to evaluation, this volume focuses mainly on program effects, adequacy of performance, and understanding of the process by which programs achieve their effects. More attention is given to ways of assessing the ultimate objective of a program—its demographic impact—than to more immediate or intermediate objectives.

In terms of the above outline, little attention is given to the evaluation of program effort, nor to possible corollary objectives of a program, such as improved child health. The determination of the appropriate quantity of effort to achieve a given demographic impact requires a long chain of calculations which works back from the ultimate objective. Chapter 3 on target setting describes part of this process—the steps and assumptions involved in going from a specified demographic objective to the number of acceptors needed to achieve this goal. The number and type of personnel needed to achieve the acceptor target require further calculations and assumptions about the number of acceptors who can be reached by specific program activities and about the quality of the various program elements. The complexity of the process connecting program inputs to demographic impact suggests that a useful type of evaluation is an

([1]) In one sense, determination of process is not an inherent part of evaluation, but attention to process makes evaluation much more valuable as an aid to administrative decision-making and often represents an important part of the scientific contribution arising from evaluative research.

examination of the internal consistency of a program with a view to determining whether the various objectives and subobjectives in terms of personnel and effects are in reasonable correspondence. Along similar lines, a type of evaluation research which can be of great practical value, though it is not treated here, is that focusing on the administration or management of a program. A framework for this type of evaluation along with a series of diagnostic questions dealing with such components as logistics, finance, control and organization has been proposed by Austin (1973).

Evaluation from the viewpoint of total demographic impact should take account of undesirable as well as desirable effects. However, relatively little attention has been given to the ways in which family planning programs might contribute to increased fertility or growth rates. In fact, it is possible that a family planning program could produce higher fertility in the short run, through treatment of subfecundity, increased survival probabilities, and the substitution for poorly understood and irregularly used program contraception for more effective private means (Reynolds, 1972: 72). Changes in age at marriage, spacing patterns, or fertility-related social customs, concurrent with and possibly related to the initiation of a family planning program, might also contribute to at least the appearance of higher fertility.

The techniques to be discussed in this overview are grouped under the three levels of effect described in Chapter 1.

(1) Techniques for assessing the extent to which the population makes use of the services offered.

(2) Techniques for assessing the effectiveness of the contraceptives used and the length of time an acceptor makes use of contraception.

(3) Techniques for assessing the effect on fertility of a family planning program. Here we distinguish between births averted analysis, based on detailed examination of the reproductive process, and analyses which relate "macro-demographic" aspects of fertility such as age-specific birth rates to program input, while taking into account or controlling for socio-economic factors which may also affect fertility.

II. RESPONSE AND USE OF SERVICES

"Acceptors of the services and supplies offered by a program are the point of departure for evaluating program efficiency and achievement of objectives. Since acceptors are people, the concept

of acceptor and a count of them might seem to be rather a straightforward matter. This, however, is not the case. Superimposed on the usual problems of keeping accurate records are conceptual difficulties of definition that arise from the potential of a duplication in count over time, space and method, and even when one method is involved, the need in many techniques for repeated acceptance of supplies.

"At the inception of the programs, it was not sufficiently appreciated that different concepts or definitions of acceptors could lead to very different counts, and hence to very different evaluation of results." (D. Nortman, discussion paper.)

As the above quotation makes clear, study of the number of acceptors from a program is the starting point for an analysis of demographic impact. If there is relatively little acceptance, it is highly unlikely that a program can have much impact. The evaluation of acceptances not only provides information on effect, but it allows evaluation of adequacy of performance through the comparison of actual acceptors with established targets, and of efficiency through the development of costs per acceptor which can be compared across countries or over time within countries.

Examination of acceptor data also provides program administrators with valuable clues for evaluating program effort and for improving program effect. Supplementary studies of why certain clinics or areas achieve better results than others can clarify the process of acceptance, and identify needed changes in program activities, personnel training, etc. The study by Lapham and Mauldin (1972) suggests that this type of approach can be expanded to an international scale in order to identify programmatic aspects associated with greater success in the recruitment of acceptors.

Chapter 2 makes clear that the concept of "acceptor" or "acceptance" must be clearly defined if useful data are to result, especially as a program matures and as additional methods of contraception are made available.

Clarification of the concept requires that attention be given to the following: which contact is to be counted? what type of contact is to be counted? which services are to be included?

The recommended procedure set forth in Chapter 2 is that only people receiving service of a specified kind for the first time be counted as acceptors. Within this rule, one may distinguish "program-acceptors", representing first-time acceptors of any method from the program, and "method-acceptors" representing first-time acceptors of a specified method. From this it follows "that an individual becomes a 'program-acceptor' at one point of time only, but that he

or she can become a 'method-acceptor' on several occasions" (Chapter 2). The receipt of services, supplies or advice (for such methods as coitus-interruptus or rhythm) with the intention to use, defines the type of contact which is counted as acceptance. In this way, use or duration of use is not generally involved in establishing acceptance. The presumption of use differs across methods, of course, and is occasionally incorporated into the data. As Nortman (discussion paper) notes:

> "Another difficulty concerning the term 'acceptor' is that for some methods, acceptance signifies intention to use—pills, condoms and other conventionals, for example—while for other methods—IUDs, sterilization or abortion—the acceptor leaves the facility as a current user of the method. The common practice is to regard as an acceptor any client accepting any method offered, thus dissociating the concept of use from the concept of acceptance. Duration of use is therefore a separate entity, calculated by various procedures for different methods. Malaysia, which offers chiefly oral contraceptives, has a category of 'provisional acceptors'. These become acceptors upon returning for additional supplies, the implication being that women who do not return did not fulfill the intention to use the pills originally supplied."

In deciding on the services to be counted, a distinction is usually made between "conception-regulating" methods and "birth-prevention" methods. The former excludes induced abortion and the recommendation of the Expert Group cited in Chapter 2 is that the term "acceptance" be restricted to contraceptive procedures, whether physiological, mechanical or chemical, including sterilization.

The basic information for data on acceptances arises from the point of contact at which services are provided. This can take the form of an individual clinic record, a coupon, or a register. The decentralized approach to record keeping has a number of implications for the quality of acceptor data:

(1) Since national totals are obtained by aggregating the data from reporting facilities, it is important that there be adequate supervision to see that the specified procedures for determining program-acceptors and method-acceptors are followed at each data-generating point.
(2) Aside from the proper classification of clients, supervision is needed to insure proper counts of the number of contacts. As Nortman notes, annual targets are often assigned to field workers, clinics and other personnel. "With pressure on field workers to meet their acceptor targets, records are known to have been inflated and inappropriate acceptors recruited." (Discussion paper.)

(3) Even with standardized record keeping and accurate reporting, multiple counting of individuals as acceptors is likely to increase with duration of the program, since those returning to a clinic or going to a different clinic will not always properly indicate their status in terms of previous contact.

To correct for duplication, it is sometimes possible to conduct special analysis of program records. Another approach is to use a survey of the target population which elicits information on the nature and frequency of contact with program facilities. In the absence of a record-keeping system at the point of contact, such surveys can also provide estimates of the number of acceptors. Such surveys do not provide unequivocal measures of the number of acceptors or the degree of duplication in acceptor data. In addition to sampling error, surveys are subject to various response and non-response errors that affect their results. In the case of acceptances, individuals who have used a program contraceptive for only a short period may not report themselves as acceptors; or contraceptives obtained from private sources may be reported as from the program. Thus the reconciliation of acceptor data from contact records with that from a survey requires judgement as to the accuracy of each system.

The basic data provided by a well organized record-keeping system are a count of the number of contacts, number of program-acceptors, and number of method-acceptors in a given time period, often monthly. It should be noted that this simple classification also provides information on the extent of return visits, which may provide a rough indication of duration of use in the absence of data needed to develop the measures of use-effectiveness described in Chapter 5. In addition to total counts, acceptors may be classified by any of the characteristics, such as age, parity, education, etc., routinely obtained on the individual record form. Chapter 2 enumerates the list of characteristics recommended as useful for evaluation and administrative purposes. Of these, information on the extent and nature of prior contraceptive use provides insight into the extent of substitution of program methods for private methods. Information on the date and nature of termination of last pregnancy indicates the degree of overlap of contraceptive use with a post-pregnancy anovulatory period, which is an important input to births-averted analysis. Along the same lines, D. Nortman (discussion paper) points out that in addition to the items shown in Chapter 2, data on duration of breast feeding and date of onset of menstruation after last birth are sometimes requested.

The simple counts of acceptors and the percentage distributions by method and by personal characteristics derivable from the basic data may in themselves prove useful to program administrators, by indicating the extent to which utilization of services is in accord with expectations. For example, a higher proportion of acceptors may be coming from older age groups or the more educated than anticipated. This in turn could lead to special studies to explain this pattern of appeal, and to changes in program procedures which attempt to redirect this pattern. Alternatively, it could lead to a reformulation of program objectives in order to accelerate acceptances among the more receptive groups.

Further guidance as to the operation of the program can be derived by cross-classification of two or more characteristics obtained in the record-keeping document. Tables 2.6 and 2.7 of Chapter 2 show the relation of method accepted to various demographic and socio-economic characteristics. Additional examples of cross-tabulations of acceptor characteristics are provided by Tables 12.5.-12.7 of Chapter 12 for Taiwan.

D. Nortman (discussion paper) points out that since age tends to be highly correlated with education and number of children, the use of an age control is desirable when examining associations with these two characteristics. When information on the referral agency or agent is obtained, tabulations of this item against age, education, parity, etc., can reveal which promotional strategies seem most successful and whether certain techniques are particularly successful with certain groups.

More refined insights into the progress of a program are possible when the basic acceptor data can be converted into a rate. This requires that appropriate denominators be available. When interest centers on the total number of acceptors, the number of currently married women of reproductive ages serves as a convenient denominator. When the data permit, further refinement may be undertaken to restrict the denominator to currently married *eligible* women. Two approaches may be distinguished: the narrower adjustment attempts to remove from the denominator acceptors in previous years, since by definition an individual can become a program acceptor only once; a broader form of adjustment takes into account additional factors which are thought to be related to acceptance, such as pregnancy status, current use of contraception from non-program sources, desire for more children, etc. An example of this latter approach appears in Freedman and Takeshita (1969: 127-130, 161-171). As noted in this source, some groups removed as ineligible in a given year, such

as the currently pregnant, may accept at a higher than average rate in subsequent years.

Though most often one must be satisfied with unrefined denominators, the effect on observed rates over time of a possibly diminishing pool of eligible persons must be kept in mind if erroneous judgements about the continued success of a program are to be avoided. Table 12.2 in Chapter 12 demonstrates clearly the difference in trend between rates based on all married women and eligible married women.

Since the number of acceptors may be cumulated across years, many programs attempt to develop rates based on acceptances since the inception of the program. A number of cautions apply to both the calculation and interpretation of such rates. Cumulative acceptance rates should not be based solely on the number of women *currently* of reproductive age, since with time many potential acceptors pass out of this age range. The more correct denominator is the number of woman-years of reproductive age since the program started, based on summing the number of women who were in the specified age range in the current year and the prior relevant years. Table 12.2 of Chapter 12 presents such rates for total IUD acceptors in Taiwan. A somewhat related analysis, which involves aging the acceptors over time, can be used with cumulative data to reveal what percentage of currently married women in a given age group have accepted in the current or a prior year. Table 12.3 of Chapter 12 presents the results of such an analysis. In addition to the problems of calculation, such rates must be interpreted carefully, particularly in comparisons across programs, since they do not refer to a specific time period but rather to the duration of the program.

Rates of acceptance can also be calculated for sub-groups of acceptors by characteristic, whenever the appropriate denominators exist, and provide further insights into the pattern of acceptance. Census or registration data as a source of denominators by a variety of characteristics are not available in many countries. In such cases, surveys can often be used to provide estimates either of numbers in a given category or of the percentage distribution of the relevant population. In the latter case, what Chapter 2 describes as the "acceptance index" can be calculated in lieu of acceptance rates. These index numbers reflect the relative over- or under-representation of each category of acceptors and serve as a useful substitute for rates. A useful caution is to insure that the classifications of characteristics used in the service statistics match those employed in the census, register or survey. If, for example, the acceptor data record education according to type of school attended, while a survey uses

level of school completed, an opportunity for the calculation of an acceptance rate or index can be lost.

Beyond the basic purposes of monitoring program progress, Chapter 2 indicates how acceptor data can be used to test models of the processes which account for the observed trend in acceptors. Though close fit of the observed trend with that expected by the model does not establish that the postulated mechanism does in fact take place, the absence of fit does tend to invalidate the model. Mathematical models which attempt to distinguish a "mass media" pattern of contraceptive adoption from an "interaction" model of adoption have recently been proposed by MacCorquodale and Pullum (1974).

In summary, analysis of acceptor data is a logical starting point for the evaluation of demographic impact. It yields a primary, and quickly available, indication of response to the program in general, to specific methods, and among specific groups. The level of acceptances can provide an early warning sign that original expectations are not being met and thus point to the need to reexamine the original assumptions of the program and/or the operation of the program machinery. Proper analysis can help identify those strategies and promotional techniques which are most effective. The main requirements for developing the appropriate data base are clear definitions of the basic concepts and the organization of an accurate record-keeping system which implements the operational measures. Such a system also generates data needed for other techniques of evaluation. Though yielding rich insights, misleading inferences can arise from over-reliance on trends in *numbers* of acceptors by characteristic, or from the use of less than optimum population denominators to monitor trends in acceptance *rates.*

III. DURATION AND EFFECTIVENESS OF CONTRACEPTIVE USE

If acceptance can be considered a necessary condition for demographic impact, it is clearly not sufficient to insure declines in fertility levels. At a minimum, factors such as the effectiveness of the contraceptive accepted, the regularity of use and the duration of use must be considered. Early awareness of this has led to considerable attention being accorded to these issues and has brought about substantial refinement in the methodology of measuring contraceptive effectiveness. These developments are the focus of Chapter 5. Three major issues are treated:

(1) The development of the life table methodology for studying the continuation rates of contraceptive use and the effectiveness of contraception at various points of time since acceptance.
(2) Since, in many programs, patterns of use following acceptance are available for only a relatively short time, the question of inferring the longer term continuation becomes important. Chapter 5 examines three techniques for extrapolating continuation beyond the period of observation.
(3) Over time, acceptors from a program may discontinue for a time and then return to the original contraceptive, or switch contraceptive methods. Chapter 5 considers extensions of the life table approach and alternatives to this technique needed to measure this more complex situation.

The techniques developed for analyzing these issues are illustrated with reference to the behavior of IUD acceptors, utilizing data from the experience in Taiwan. The methodological treatment, however, is general enough that with minor modifications it is applicable to other settings as well as for studying the post-acceptance behavior of acceptors of other methods of family planning.

The following analysis of the conceptual and measurement aspects of these techniques are, with the exception of minor editorial changes, taken from the discussion paper prepared by A. K. Jain. At a few spots, additional material supplied by the editors has been inserted and this is set off by brackets to distinguish it from Dr. Jain's review.

LIFE TABLE APPROACH

"The life table approach for studying the use-effectiveness of the IUD was developed independently by Potter and Tietze during 1966-1967. Since then, it has been widely used around the world in clinical as well as in field settings. With some modifications, the life table approach has also been used to study the use-effectiveness of the pill. This approach has distinct advantages over the Pearl Index, which has been commonly used to measure the effectiveness of contraception. It provides not only an index of effectiveness, but also indices of the continuation of a method and the relative intensities of the several causes for discontinuing the use of a method. The values of these indices are provided for monthly or broader time intervals following the adoption of a method, thus making it possible to study the changes in the risks of termination as a function of time elapsed since the adoption. The approach has an additional advantage in that it incorporates into the analysis the incomplete histories,

which are truncated due to a fixed cutoff date, as well as the complete histories.

"Two types of termination rates, 'net' and 'gross', are identified. Net termination rates due to a specific cause, for example pregnancy, are calculated in such a way that they allow for the presence of other competing risks of termination, such as expulsion and removal. Thus, the net termination rates for a specific cause, say pregnancy, are reduced because some women expel the device or have it removed before they have had a chance to conceive accidentally. The net termination rates for each specific cause are additive and the sum of these rates equals the total termination rate for all causes. However, net termination rates do not adequately measure the force of termination risk specific to a particular cause, independent of the competition from other causes. Gross termination rates measure this force; they are slightly higher than the net rates and are not additive. They are recommended for use in comparing the force of termination risks for different samples or for different sub-groups of the same sample. The gross rates are also useful for studying the changes in intensity of particular termination risks with duration from acceptance.

DATA REQUIREMENTS AND MEASUREMENT PROBLEMS

"Two types of information are required for application of the life table technique to the study of use-effectiveness of contraception: (1) type of termination and (2) ordinal months of use. The required data can be obtained either through prospective clinical studies or retrospectively through field surveys. In the case of IUDs, the authors have considered five major categories: users, pregnancy, expulsion, removal and lost-to-follow-up (LFU). Acceptors are divided into these five categories according to their status with respect to the use of the IUD at the cutoff date. Sometimes these categories are further subdivided, depending on the specificity required. For example, the pregnancy category can be divided into: (a) pregnancy with the IUD *in situ* and (b) pregnancy with the status of the IUD undetermined. The second category includes those pregnancies which occur following an unnoticed expulsion of the IUD. Computing the ordinal months of use requires accurate information about the dates (day, month and year) of adoption and termination of a method. The ease of accurately ascertaining these dates varies by type of method and type of study, clinic or retrospective.

"Acceptors who are classified as users or LFU at the cutoff date create problems in the calculation of ordinal months of use. For the LFU category, the recommended procedure is to consider only the

period up to the last contact with the acceptor, because at this contact the IUD was known to be *in situ.* It is usually assumed that the acceptors' experiences with the IUD between the last contact and the cutoff date are similar to those of subjects who have been successfully contacted during this period. The bias will be minimal if the proportion of acceptors classified as LFU is low and/or the relevant assumption is correct. A reasonable assurance for making such an assumption is provided by comparing the characteristics of those classified as LFU with the corresponding characteristics of those for whom complete information is available. Total termination rates will be underestimated or overestimated, depending upon whether the LFU cases have higher or lower risks of termination than those who have been contacted successfully. Under these circumstances, a range of termination rates could be obtained under two extreme assumptions: that all of the LFU cases terminated the use of the IUD immediately after the last contact; or that all LFU cases continued using the IUD until the cutoff date. The latitude of this range will depend upon the proportion of total acceptors classified as lost-to-follow-up. The range will be narrow if this proportion is small and it will increase with the increase in the proportion of the LFU cases.

"In studies based on clinic data, acceptors classified as users can be treated in two ways. Tietze has recommended that they be given credit up to the cutoff date, whereas Potter gives them credit only up to the date of contact. This distinction has never been brought out in the literature and deserves some consideration. The users category consists of two groups:

(1) those for whom the scheduled follow-up visit at the cutoff date is overdue by up to 90 days, and
(2) those for whom the follow-up visit is scheduled after the cutoff date.

These two categories are treated separately in the following discussion.

"Following acceptance from a clinic, a uniform schedule for follow-up visits is prescribed. This schedule usually recommends follow-up contacts at 1, 3, 6 and 12 months following insertion, and every year thereafter. Following the example given by Potter and Avery, let us consider the cutoff date as December 31, 1971. Let us also consider those cases which were last contacted in 1970. Among these cases, those who were contacted prior to September 30, 1970, will be classified as lost-to-follow-up. For these cases, the follow-up visit was due prior to September 30, 1971, and since the period

between the scheduled follow-up visit and the cutoff date is more than 90 days, they will be classified as lost-to-follow-up, and credit for IUD use will be given to their last contact prior to September 30, 1970. However, those who were contacted during October-December, 1970 had a follow-up visit due during October-December, 1971. Since the period between the scheduled follow-up visit and the cutoff date is less than 90 days, these cases will be classified as users. According to the Potter technique, the months of IUD use for these cases will be calculated up to the last contact during October-December, 1970. However, according to the Tietze technique, the months of IUD use for these cases will be calculated up to December 31, 1971. The difference between the two techniques could be as much as 15 months. Depending upon the number of such cases, the effect on the termination rate could be high. In the example used by Potter and Avery, we do not know whether or not there were any such cases.

"The second category of users includes those for whom the next follow-up visit is scheduled after the cutoff date. For example, for those acceptors who were last contacted during 1971, the next follow-up visit will be scheduled during 1972, which will be after the cutoff date of December 31, 1971. For these cases, the credit given for IUD use according to the Potter methodology will be different from that given according to the Tietze methodology. The difference between the two could vary between one and twelve months. In their example, Potter and Avery have shown that the effect of the differences in computing ordinal months according to these two methods on the termination rates is negligible. This may not be true under all circumstances. The net effect will depend upon the number of such cases as well as upon the distribution of their last contact during 1971. For example, if for most acceptors the last contact was made during the first half of the year, then the effect on the termination rates will be greater than if the last contact was made close to the cutoff date.

"A third category of problem cases includes those acceptors who do not return for any follow-up visit. These cases are divided among users and lost-to-follow-up, depending on whether their scheduled follow-up visits are overdue by up to 90 days or by more than 90 days. According to the Tietze methodology, the LFU cases in this category are given one month's credit for IUD use and those classified as users are given credit up to the cutoff date, that is, ranging from one to four months. According to the Potter methodology, however, all these cases are probably excluded from the analysis. The effect on termination rates would again depend upon

the number of such cases and will only affect the termination rates in the initial periods.

"As an illustration, the effect on the total termination rates of the above-mentioned differences in treating users and lost-to-follow-up cases is shown in Table 13.1. The total termination rates of Copper-T model 200 (hereafter referred to as TCu-200) in the United States are computed under three different assumptions. These data are taken from a multicenter prospective ongoing study of the TCu-200 in the United States being conducted under the auspices of the Population Council. December 31, 1973 is taken as the cutoff date for this analysis.

Table 13.1. *Total cumulative termination rates of TCu-200 per 100 woman-years by ordinal months of use under three different assumptions: United States, cutoff date, December 31, 1973*

Ordinal months	*I*	*II*	*III*
1	3.5	3.8	3.7
3	8.7	9.5	8.1
6	14.7	16.3	13.0
12	25.6	28.7	21.2
18	34.9	39.6	27.7
24	42.2	48.4	32.7
30	50.7	59.2	37.6
36	60.1	69.8	42.0

I. Ordinal months for lost-to-follow-up cases are calculated up to the last clinical contact and for users are calculated up to the cutoff date. Women with no contact since insertion are divided among users and lost-to-follow-up categories, depending upon whether their visit was overdue by up to 90 days or more than 90 days (cf. Tietze).

II. Ordinal months for lost-to-follow-up cases and users are calculated up to the last contact. Women with no contact since insertion are excluded (cf. Potter).

III. Ordinal months for lost-to-follow-up and users are calculated up to the cutoff date (lower limit).

"Under the first assumption, cases classified as lost-to-follow-up are credited for TCu-200 use up to the last contact, and those classified as users are credited up to the cutoff date. Those with no contact since insertion are classified as users or lost-to-follow-up, depending on whether or not their scheduled follow-up visit was overdue by 90 days. This scheme corresponds to Tietze's recommended procedure. Under the second assumption, cases classified as either users or lost-to-follow-up are credited only up to the last contact, and cases with no contact since insertion are excluded. This scheme closely corresponds to Potter's methodology. In the third assumption, the

ordinal months for lost-to-follow-up and users are calculated up to the cutoff dates. This assumption provides a lower limit for termination rates which can be used as a reference point for the effect of other assumptions.

"The results in Table 13.1 do not give the full range of possible outcomes. For example, if those lost to follow-up were treated as *terminators* at the point of last contact while users are calculated up to last contact, the cumulative termination rates at 24 months would be 62 per 100 woman-years, 14 per 100 more than the largest rate shown in Table 13.1 at that duration.

"The results in Table 13.1 indicate that the differences in termination rates under different assumptions increase with the duration of use. Let us consider the magnitude of the termination rates under the first two assumptions. According to the Potter methodology, 48.4 per cent of women terminated the use of the device at the end of two years, as compared with 42.2 per cent according to the Tietze methodology. The difference between the two is about 6 percentage points, increases to about 10 percentage points at three years and is likely to increase further beyond three years of use. The difference mainly reflects the effect of differences in the credit given to users for IUD use. Whether credit for IUD use should be given up to the last contact or up to the cutoff date remains a question which needs some further consideration.

"It is usually assumed that women who terminate the use of the device are more likely to come to the clinic than those who are happy with the device. The latter group is likely to forget the scheduled follow-up visit, though in fact they are likely to be using the device. Using this rationale, credit is given for IUD use among users up to the cutoff date. However, if this assumption is correct, then it should also apply equally to the LFU cases. If we do just that, then the termination rate drops to 32.7 per cent at the end of two years. Under the circumstances, one would be tempted to conclude that the termination rate lies between 32.7 per cent and 42.2 per cent. However, this will reflect an optimistic picture of the real situation.

"A better alternative seems to be to treat both the LFU cases and the users in the same way, and to give credit for IUD use up to the last contact for acceptors in both these categories. This will, no doubt, yield a conservative estimate of the termination rates. However, one would be reasonably assured of its magnitude, because there is no reason to believe that the termination risk between last contact and the cutoff date among users and LFU cases will be higher than the termination risk among those who are successfully contacted near the cutoff date."

[Another implication of the foregoing analysis is that differences in continuation rates across clinical studies can be a function of the specified schedule of revisits, in addition to the way credit is allocated to users and LFU cases. Since revisit schedules are not uniform across settings, this suggests that attempts be made to interview all acceptors within a narrow span of time, so that their status at the cutoff date will be accurately known.]

INTERPRETATION

"The interpretation of results obtained by applying the life table methodology is quite simple. Potter and Avery provide some guidelines by using the life table technique to analyze the first segment experience of IUD acceptors in Taichung City in Taiwan. These results are shown for a period of eight years. In order to be brief, the authors have shown these results at six-month intervals in Table 5.1 of Chapter 5. These results adequately explain the experience of IUD use for a synthetic cohort of women. The interpretation of these results includes two parts: a description of the situation at the end of the observation period, and understanding the trend. An example follows:

"In Taichung, one out of five women who accepts an IUD continues to use it for at least eight years: that is, one fifth of the acceptors get full protection from the IUD against the risk of pregnancy for as long as eight years (2). The remaining four-fifths do not get complete protection for a full eight-year period, because some of them become pregnant during this period, some expel the device and others have it removed. The relative magnitude of these three risks is shown by the cumulative termination rates, which are also obtained by using the life table technique. However, these rates are not shown in Table 5.1. Earlier results from this survey have indicated that removals account for a majority of terminations, followed by expulsions, and that terminations due to pregnancy are few. An example of cumulative termination rates by type of termination and characteristic of acceptor is given by Table 12.15 of Chapter 12.

"Potter and Avery have used six-month conditional probabilities for describing the relationship between termination risks and the duration of time that has elapsed since insertion. These conditional

(2) This proportion is based on a synthetic cohort in which women observed for a shorter period contribute to the corresponding probabilities of termination. Table 12.14 (Chapter 12) shows the proportion of women observed for 96 months who have the original IUD. In this case the two proportions are very close, though this need not be the case, especially when later acceptors differ appreciably from earlier ones in their probability of termination.

probabilities of termination for all causes decline with the increase in time elapsed since insertion. One might be tempted to interpret this decline as indicating that the risks associated with IUDs decrease with the increase in the time which has elapsed since insertion. However, such an interpretation will not be accurate because it assumes that women are homogeneous with respect to the risk of termination. Potter and Avery have provided a more plausible interpretation by assuming heterogeneity among women with respect to the risk of termination. According to this assumption, women with a higher risk will terminate earlier, leaving behind a group of less heterogeneous women. Thus, the heterogeneity among women, not the risk of termination among individual women, will decrease with the increase in the amount of time elapsed since insertion. This interpretation is also supported by the fact that the age and parity differentials in six-month probabilities of termination also decrease over time. For example, among those who accepted the IUD for limiting their family size, the probability of continuation during the first two years following insertion increases from 0.328 to 0.652 with the increase in age at insertion (Table 5.2, Chapter 5). However, during the period 6-8 years following insertion, these probabilities are close to 0.8 for all age groups. Nevertheless, some decrease in the risk of termination for individual women cannot be completely ruled out, especially for accidental pregnancies. It is quite unlikely that the fecundability of individual women will remain constant for as long as eight years. With the decrease in fecundability, the risk of accidental pregnancy will also decline. Thus, in addition to heterogeneity, decreasing fecundability will account for the observed decline in six-month conditional probabilities of IUD termination due to pregnancy.

EXTRAPOLATION BEYOND THE PERIOD OF OBSERVATION

"The results obtained by using the life table technique are not sufficient to estimate the total protection provided by the first segment of IUD use. This information is required for estimating the number of births averted by IUD use. However, the results obtained by using the life table technique provide a basis for extrapolating beyond the period of observation. Potter and Avery extensively review the applicability of two models for extrapolation. The first model was suggested by Mauldin *et al.* (1967) and is based on a single negative exponential function. The second was suggested by Liu *et al.* (1972) and is a mixture of the function of two negative exponentials. Potter and Avery show that both these models provide conservative estimates

of the average retention period of the IUD, especially when the observation period is short. The observed average retention period during 96 months in the Taichung sample is 37.5 months. For the experience of the first year, the average retention period during 96 months was estimated at 27.5 months based on the Mauldin *et al.* technique and 28.7 months based on the Liu *et al.* technique (Table 5.3, Chapter 5). On the other hand, if the experience of 96 months is utilized, the estimated average retention period becomes 36.1 months based on the Mauldin *et al.* method and 38.8 months based on the Liu *et al.* method. The second model, in fact, over-estimates the retention period at longer durations. The estimates of average retention period by both techniques are affected by the period of observation.

"Potter and Avery have suggested another model to minimize the effect of the bias introduced by length of observation. Their model is based on the assumption of a fixed risk of discontinuation for each acceptor and of heterogeneity among acceptors with respect to this risk. This risk among acceptors is assumed to vary according to a Pearsonian Type III curve. The proposed model has three advantages over the Mauldin *et al.* and Liu *et al.* techniques:

(1) the estimates of the parameters of the proposed model are far less sensitive to the length of observation,
(2) the predictions of average retention period are fairly adequate, and
(3) the overall fit of the model is better.

A distinct disadvantage of the proposed model is that the estimated value of the average retention period is defined only for finite values of the maximum duration of IUD use. Recognizing this limitation, Potter and Avery have indicated that one must adopt some maximum duration—for instance, the duration that would bring ages to an average of 45 years. In the Taichung sample this maximum duration would be about 165 months, calculated from the age distribution given in the chapter. Using 96 months of experience, the estimated average period of retention during 165 months will be 45.7 months according to the Mauldin *et al.* technique, 47.8 months according to the Liu *et al.* technique and 49.3 months according to the Potter and Avery technique. These three estimates are quite close. However, if only the experience of the first two years were available, then the average retention period during 165 months estimated by using the three techniques would be 32.5, 33.4 and 49.7 months respectively.

This means that for shorter observation periods, the technique suggested by Potter and Avery will probably predict the average

retention period more accurately than the technique suggested by either Mauldin *et al.* or by Liu *et al.* Since in most studies information is available only for a short period of observation, the technique suggested by Potter and Avery could be used with greater confidence."

EXTENSIONS OF LIFE TABLE PROCEDURE AND ALTERNATIVES

[The life table procedure is most directly applicable to analyzing a single segment of contraceptive use and, among contraceptives, it is most efficiently applied to IUDs where the problem of irregular practice does not arise and hence the difficulty of definition of first segment length is minimized. With the passage of time from acceptance, the behavior of acceptors with respect to contraception and fertility grows more complex: in the case of IUDs there is a high frequency of reinsertions; with oral or conventional methods, there may be periods of non-use; upon termination of one method, an acceptor may switch to another method or cease to use contraception; pregnancies followed by abortion or live births can intervene. Undue focus on the length of the first segment may lead to pessimistic inferences about the success of a program. For example, administrators noting that less than half the IUD acceptors maintain the device for two years may assume that those terminating give up contraception entirely and resume a high rate of child-bearing. However, Freedman and Takeshita (1969: 347) have shown that if one focuses on the acceptors rather than the duration of the first device, the picture is quite different. At 30 or more months from insertion, about half the acceptors had either the original or a reinserted IUD *in situ*, and another quarter were currently using some other contraceptive (including sterilization). Only 14 per cent had experienced a live birth. A similar analysis of contraceptive-pregnancy status by duration is given in Table 12.14 of Chapter 12.

The results of applying the life table technique are not adequate for capturing the dynamic shifts in contraceptive behavior which can take place following acceptance nor for monitoring the fertility experiences by a cohort of acceptors.

In Chapter 5, three extensions of the life table procedure which are designed to handle certain aspects of this complex picture are described. Reinsertion rates can be developed for monitoring the tendency of IUD acceptors to return to this method at various times after terminating the first device; multiple segment analysis which ignores brief periods of non-use and is applicable to oral and conventional methods of contraception as well as the IUD; extended use-effectiveness, which focuses on the time during which a woman avoids

pregnancy or a live birth, and ignores the fact that there may be switching from one method to another, or periods of non-use.

None of these methods, however, nor the recording of the status of acceptors at various points in time since acceptance or termination exemplified in the Freedman and Takeshita analysis (cited above) captures the detailed patterns of change in methods or periods of non-use. Toward this end, the authors of Chapter 5 propose a technique based on Markov Chains.]

"This approach provides a series of matrices, one matrix for each interval of a specified duration. Each matrix provides probabilities of remaining in the original state during a specified period as well as transitional probabilities of moving from one state to another during the same specified period. The number and type of states are flexible and can be selected according to the purpose of the study. Many interesting results can be obtained by using this approach. By using the original distribution of acceptors among the specified states and these matrices, one could obtain the expected distribution of acceptors among different states at the end of the study period. For example, at the end of a five-year period following acceptance of the IUD, the proportion using other forms of contraception, the proportion that have been sterilized, the proportion not using any method and even the proportion that would have become pregnant and would have given birth can be estimated. By studying the transitional probabilities, one could understand the mechanism by which the end result is achieved. One could also study the relative attracting power of each state.

"Potter and Avery have considered four states: IUD, other contraception, no contraception and sterilization. They used data from Taichung to illustrate the application of this approach, and have shown that the long-term acceptability of the IUD (including the Ota ring) is high enough for it to recoup its losses to other methods or to non-use of contraception by subsequent reacceptance. This conclusion should not be treated as universally true, since the acceptability of a particular method will depend upon many factors which may vary from one situation to another. The authors do not make any suggestion that their conclusion will be true in other situations. The main point is that by using this approach, the relative long-term as well as the short-term acceptability of different methods in a given situation can be studied.

"The data required for applying this approach are not usually available, because women are not followed after they terminate the use of a particular method. This type of data, however, needs to be collected, and one hardly needs any justification for doing so. The

suggested approach provides great promise for future studies directed toward the understanding of contraceptive behavior.

COMPARABILITY OF RESULTS

"One may use the life table procedure to measure use-effectiveness to:

(1) describe the experiences of a cohort of couples with one method in a specific setting,
(2) compare the performance of a method in different settings, and
(3) compare the performances of different methods in the same or different settings.

"In making comparisons across settings or across methods, attention should be given to the many factors affecting the measured effectiveness of a contraceptive or continuation rates, other than those explicitly taken into account in the life-table analysis. For example, a post-partum program and a program in which the open interval of acceptors is longer may be expected to show differences in pregnancy rates because the two populations will differ in the amount of overlap of contraceptive use with the post-partum anovulatory period. Programs which differ in the proportions of those accepting for the purpose of spacing births rather than limiting family size can be expected to differ in termination rates because of the shorter average duration of use of the former group. Differences in the composition of two study populations on socioeconomic and demographic characteristics such as age, parity and education can also affect termination rates. Where the data permit, the life table analysis can be carried out for various sub-groups of the populations, thus enhancing comparability. In addition to the foregoing factors, attention must be paid to possible differences in the procedures for computing ordinal months of use for those classified as users or lost-to-follow-up, and the revisit schedules of the clinics, as described earlier in this section."

SUMMARY

[The life table procedure provides a highly refined technique for measuring the duration of use after acceptance and for determining the effectiveness of the contraceptive. Despite its power, comparisons must be made cautiously since differences can arise from the method employed to classify users and those lost-to-follow-up, and from differences in the demographic and socioeconomic char-

acteristics of the target populations. Since the probabilities of termination change with time from acceptance, and given the need for some estimate of the average length of use based on a limited period of observation, methods of extrapolation are needed. Chapter 5 indicates some useful approaches to this problem. The life table technique has only limited ability to monitor the complex patterns of contraceptive use and non-use that develop over time. Extensions of the basic approach provide measures that capture part of the process, but fail to reflect the full dynamics. The application of Markov Chains to this problem, described in Chapter 5, seems most promising and should be tested in different settings to explore its full potential.]

IV. PREVALENCE OF CONTRACEPTIVE USE

The number of acceptors each year and their continuation rates jointly determine the number of couples using contraception at any point in time and allow one to calculate the proportion of women of reproductive age currently using a program method. Measures of prevalence are therefore useful indicators of program success in that they reveal the potential demographic impact, as well as providing an estimate of the size of the residual group which needs to be reached.

Data on the proportion of couples currently using a contraceptive can of course be estimated from surveys, though it is often desirable to have this measure on a more regular basis. Toward this end, estimates of continuation rates appropriate to each contraceptive method and period of acceptance are sometimes employed. Tables 12.12 and 12.13 illustrate the calculations employed in Taiwan for assessing the prevalence of IUDs. When methods other than the IUD are involved, one must often rely on other than life table procedures to estimate continuation rates. Chapter 12 describes two approximations used in Taiwan for pills and condoms.

Estimates of prevalence have also been derived from service statistics using the couple-years of protection (CYP) approach (Wishik and Chen, 1973). This technique is also used to produce a measure of program achievement as described in Chapter 2. For this purpose, CYP attempts to measure the joint impact of all program methods, taking into account the length of time a couple is protected by the method adopted. It is a weighted sum of the methods distributed each month, with the weights representing the duration of time for which the method is presumed effective. The annual CYP is obtained

as the sum of the monthly totals. In estimating prevalence of use during a given period by the CYP technique, one must include protection resulting from methods distributed during that period as well as protection resulting from past distribution. Thus, for methods such as the IUD or sterilization, which can afford long-term protection, distributions of prior years must be carried forward with weights specific to both method and time of distribution. For IUDs, the weights will reflect the probabilities of continuation from the time of insertion to the period in question, while for sterilization, they will be based on the probabilities of survival and maintenance of a fertile union.

V. BIRTHS AVERTED ANALYSIS

Though measurements based on number of acceptors and use-effectiveness through life table procedures provide valuable information on essential elements of a family planning program, they are still at some remove from the question of the program's impact on fertility. As Chapter 1 notes, two broad approaches to ascertaining demographic impact may be distinguished. The first, births averted analysis, focuses on the components of the reproductive process and seeks to determine the effect of a segment of program contraceptive use on this process. Four chapters in this volume, Chapters 3, 4, 6 and 10, have relevance for this approach. The following discussion of these chapters by R. G. Potter brings out the essential features of births averted analysis, the types of data required and the utility as well as the limitations of the various methodologies designed to capture this aspect of a family planning program.

"In broadest perspective, the measurement of births averted is quantifying the impact of regulatory efforts upon fertility. To affect childbearing and bring it into closer conformity with couples' spacing and family size preferences, control practices such as contraception, induced abortion and sterilization must operate through a set of factors that the French demographer, Louis Henry, has called 'natural fertility'. This system encompasses strictly physiological factors such as fecundability, intra-uterine mortality, gestation length, the temporary anovulatory periods following delivery and primary and secondary sterility; but it also includes social mores governing entrance into and out of marriage, breast-feeding practices, extent of sexual contacts outside of marriage, work-imposed separation of the spouses and other similar factors. Though social in nature, these customs, for being imposed upon couples and not at their disposal to manipulate according to particular size or spacing goals, are properly viewed as

part of the natural fertility system and indeed help to explain why the level of natural fertility is so variable from group to group.

"Henri Leridon's balanced review of the biostatistics of natural fertility, in Chapter 4, provides essential background for three of the chapters in this volume. First is Chapter 3 in which Ross surveys procedures for setting acceptor targets that will realize predesignated demographic goals. These procedures necessarily deal with births averted. Second is Chapter 6 containing Wolfers' critique of methods for estimating births averted among cohorts of program acceptors. Third is Menken's Chapter 10 covering simulation studies as an aid for assessing the effects of a family planning program. Especially in the more elaborate Monte Carlo models, detailed assumptions about natural fertility as well as family planning practices have to be made.

"What Leridon calls 'macro-demographic' aspects of fertility—the most overt and directly measurable features of fertility—include age-specific birth rates, marital age-specific birth rates, distributions of completed family size, and mean birth intervals. These indices reflect the interaction of a host of 'microphysiological' factors that are conveniently resolved into four main elements: 'natural fecundability', risks of intra-uterine mortality, lengths of nonsusceptible periods, and risks of sterility.

"A couple's natural fecundability is the woman's monthly probability of conception when not contracepting and fecundable (i.e., not pregnant, temporarily anovulatory after a pregnancy, or sterile). Intra-uterine mortality includes both early pregnancy losses (spontaneous abortion) and later ones (stillbirths) as well as the very early fetal losses that go unrecognized by the woman and therefore downwardly bias estimates of natural fecundability. Nonsusceptible periods include both the duration of pregnancy and its ensuing span of temporary anovulation. The lengths of nonsusceptible periods are conditional upon the outcome of the pregnancy and the survival of the child, being typically longer after childbirth than after a fetal loss. Sterility is primary if the woman can produce no children; secondary if the condition follows one or more births.

"The above scheme is readily extended to represent the action of fertility control. Contraception acts by reducing natural fecundability to a lower level, denoted 'residual fecundability', thereby tending to hold the woman in the fecundable state longer than would otherwise be the case. (The 'effectiveness' of a contraceptive is technically defined as its proportionate reduction of natural fecundability.) The practice of induced abortion raises the rates of intra-uterine mortality. Sterilization usually advances the onset of secondary sterility.

"A number of points in Leridon's review have particular relevance

for estimating net births averted per acceptor of contraception. Such estimates involve two subanalyses: first, to estimate the 'potential fertility' of the acceptors if they had not accepted the program methods they did and second, to gauge the spans of 'effective use' of the chosen methods, during which time potential fertility is being foregone.

"Estimates of potential fertility most often are drawn from data of the macrodemographic type: either age-specific rates or, better, marital age-specific birth rates, or else mean birth intervals. Microphysiological data, together with appropriate information concerning contraceptive practice, are usually utilized in assessing length of effective use.

"The contrasting ranges of age-specific and marital age-specific birth rates indicate the important roles played by age at marriage, divorce and widowhood. The two measures are not at all interchangeable. Accordingly, marital status constitutes an essential control when measuring potential fertility. Under conditions of natural fertility, both marital age-specific and age-specific birth rates decline with advancing age after about age 25. Thus, age too is an indispensable control when considering potential fertility. Variation in sets of marital age-specific birth rates is such as to indicate wide variation from group to group in natural fertility. Hence, even when an alternative to program contraception is being assumed absent, there is no one standard set of natural birth rates to invoke as the acceptors' potential fertility. Rather, the acceptors' preprogram fertility with suitable adjustments for aging must be used or else the fertility of a carefully matched sample be employed, though, as noted later, each of them presents some difficulties.

"Under conditions of natural fertility, among a single group of women married at about the same age, and whose marriages remain intact until the end of childbearing, wide variation in completed family size is found. Contributing to this variation are both differing ages at last birth and differing mean birth intervals. Women of high final parity tend to have their last child late and to evince shorter than average birth intervals.

"There is no simple relationship between the reciprocal of marital age-specific birth rates and mean birth intervals. From the decline in marital age-specific birth rates with advancing age and from the close correlation between a mother's age and her parity, one would expect parity-specific birth intervals to lengthen with increasing parity. As Leridon demonstrates, this expected result does not obtain. Only when one compares parity-specific mean birth intervals among women sharing a common completed family size does the expected decline

emerge. Without control on final parity, the expected trend of mean birth intervals is distorted by differential selection for shorter than average intervals associated with large completed family sizes.

"A similar phenomenon is encountered when dealing with mean birth intervals classified by age of mother. Variation in length of birth intervals combined with the finite length of the reproductive period produces 'truncation biases'. One may classify birth intervals by age of mother at the start of the interval or by her age at its end. Let us call the two results 'prospective' and 'retrospective' birth intervals, respectively. Mean prospective birth intervals tend to be biased downwards at older ages, since if an interval begins with the woman at a relatively advanced age, its chances of being completed before the onset of secondary sterility increase the shorter it is. Correspondingly, mean retrospective intervals tend to be biased downwards at younger ages, since a relatively young age at the end of a birth interval is associated with a selection both toward young marriage age and short birth intervals. Thus, mean prospective birth intervals generate quite different values (except at intermediate ages) from mean retrospective ones.

"Whether by chance factors or by systematic fecundity differences, some women bear children at shorter intervals than others. Within a given marriage duration, those who bear children more frequently average more birth intervals than those who bear infrequently. Hence a simple averaging of birth intervals—called by Wolfers a 'mean birth interval (births)'—yields lower values than a mean birth interval in which each woman contributes at most one birth interval. The latter Wolfers denotes as a 'mean birth interval (women).' Clearly, then, there are a variety of mean birth intervals that may be computed and when using birth interval data for purposes of estimating potential fertility it behooves one to have a clear rationale for choosing one index over others.

"Leridon's consideration of microphysiological factors makes clear the multiple bases for fecundity differences among women. The large variation in completed family size found under conditions of natural fertility is attributable in part to differing ages at marriage and at onset of secondary sterility. Women of like age may vary with respect to natural fecundability, proneness to intra-uterine mortality, and to expected lengths of postpartum anovulation.

"Leridon reviews both direct and indirect evidence indicating that given prolonged breastfeeding, with effects perhaps intensified by an inadequate diet, the period of postpartum anovulation may average as long as a year. Two important implications follow from this sizeable component of birth intervals. First of all, the intervals

from marriage to first birth may be expected to average many months shorter than subsequent birth intervals, since the latter include a period of postpartum anovulation following the initial birth, while the former intervals, starting with marriage, do not. Secondly, under a regime of natural fertility, a nonsterile woman is under risk of conception only a fraction of the time. During much of her reproductive period, she is protected either by pregnancy or postpartum anovulation. One way in which the effect of contraception is reduced is by prolonging stays in the fecundable state, thereby forfeiting some of the natural protection time conferred by pregnancy and postpartum anovulation.

"The consideration of microphysiological factors also helps to explain the regularly observed decline in natural fertility rates with advancing age. As a rough generalization, about half the decline from age 25 onwards is attributable to the rising proportion of secondarily sterile women. The rest of the decline is some combination of diminishing natural fecundability, rising risk of intra-uterine mortality, and lengthening postpartum anovulation. During the early years of the reproductive period, starting with menarche, the trend of average fecundity is upwards, but the details are poorly measured for lack of data. At these early ages, trends in birth rates and birth intervals are complicated by rapid changes in marriage-to-first-birth intervals as a fraction of all birth intervals.

"Secondary sterility is primarily a function of age. A small additional risk is thought to associate with childbirth. No important parity effect *per se* has been detected. If a woman has just had a birth, then her risk of being secondarily sterile is less than that of a woman taken at random, the latter having a small chance of being primarily sterile as well as having been subject to the risk of secondary sterility since her last birth, which might have been some time ago. The degree to which recently parturient women have lower rates of secondary sterility than their age mates taken at random increases rapidly after age 35, when proportions sterile have reached significant magnitudes.

"What are the implications of the above generalizations for the difference in average fecundity between acceptors and nonacceptors of similar age? The large chance component involved in family-building histories makes it easy to exaggerate the relative potential fertility of acceptors. In many developing nations women are more likely to accept contraception when they reach or pass desired family size than before its attainment. Acceptance at an early age selects for an interest in spacing, for fewer children desired than average, for an earlier than average marriage age, but also for shorter than

average birth intervals. It is tempting to project this rate of childbearing into the future (perhaps adjusted for aging) as an estimate of potential fertility. Yet to the extent that these relatively short birth intervals reflect chance factors rather than superior fecundity, their extrapolation will tend to overstate future fertility. Thus, the fact that acceptors have shorter birth intervals, or higher birth rates, in the past than nonacceptors may not mean that they have significantly higher fecundity. The exception may be at advanced ages. If acceptance tends to follow a birth closely, then among older acceptors—over 35 years, say—one might expect to find appreciably lower rates of secondary sterility and on that account a higher potential fecundity than among coeval nonacceptors.

"Attention now turns to Wolfers' Chapter 6 and procedures for estimating net births averted per acceptance of a program method. For the most part, that method is taken as the IUD, though the concepts treated are general enough to fit most other contraceptives as well. Contraception averts births indirectly by keeping women in the fecundable state longer than would otherwise be the case, thereby interrupting the cycle of exposure to conception risk, pregnancy, anovulation and back to exposure. If the interruption is long enough to equal the marriage duration necessary for a potential birth in the absence of the chosen method, then it is interpreted that a birth has been averted.

"As already noted, this approach involves two basic calculations, both considered in detail in Wolfers' chapter: first is estimating lengths of 'effective usage', and second is gauging the potential fertility to be expected, had the program method not been accepted. The latter is expressed as a marriage duration per potential birth. Dividing this duration into woman-years of effective use yields births averted. Typically the analysis proceeds within age classes.

"To prolong a woman's stay in the fecundable state, a contraceptive must, on average, reduce natural fecundability during its practice. A contraceptive does no good if not practiced because then it is not lowering fecundability. Likewise, it accomplishes nothing if practiced by a woman during postpartum anovulation or when she is already sterile. In either case, fecundability remains at zero. Furthermore, x months of practice that ends in accidental pregnancy is not conferring the same degree of protection as x months of practice without a failure.

"Gross wearing time of an IUD is most conveniently assessed by a life table analysis of continuation, a topic addressed in Chapter 5. Dislike of the method, physical intolerance of it, or deliberate interruption in order to have a child, are among the motives for discon-

tinuing; divorce, widowhood or death of the woman herself are additional reasons. To estimate average wearing time per device inserted, it becomes necessary to extrapolate the proportions retaining their device beyond the period of observation. Techniques for doing this are also covered in Chapter 5. Total wearing time is converted into effective wearing time—that is, time when practice of the method does some good—by means of a series of adjustments: a deduction for overlap with postpartum anovulation, another to allow for the fraction of women sterile at the time of insertion or becoming sterile during usage, and finally a penalty per accidental pregnancy.

"Length of overlap between wearing time and postpartum anovulation varies greatly from group to group. Given a postpartum program and lengthy lactation, for example, then only a minority of acceptors would have no overlap. Direct data on the distribution of overlap is unlikely. The next best situation is to have the distribution of 'open intervals' between last birth and acceptance, combined with an estimate of mean postpartum anovulation. The latter estimate can be converted into a probability distribution of postpartum anovulatory lengths by means of the Pascal distribution; and then assuming statistical independence between open interval and postpartum anovulatory lengths, a probability distribution of overlaps can be derived. When still less data are available, one is forced into crude, seriously biased allowances for overlap.

"The adjustment for overlap with secondary sterility rests on the premise that since last birth acceptors have been subject to annual (or monthly) hazards of secondary sterility appropriate to their age and that these risks continue to operate during retention of the device. Wolfers starts the period of sterility exposure nine months earlier at the time the last birth is conceived.

"It is recommended that for each accidental pregnancy, one deducts from woman-months of effective use the expected number of fecundable months required per conception in the absence of the program method. When the rate of accidental pregnancy is low, as with the IUD and oral contraception, this adjustment proves of secondary importance. Greater effect may be expected in the case of coitus-related methods when irregular practice is augmenting failure rates.

"A second, basic calculation is the mean duration required per potential birth by fecund women belonging to a specified age class. There exist two main sources of data. First are mean birth intervals. By definition, only fecund women have birth intervals so that a direct adjustment for secondary sterility becomes unnecessary. Wolfers recommends use of mean prospective intervals (births). Being

averaged are the birth intervals of women belonging to a certain age class at the start of the intervals. This choice allows for the fact that those who bear children more frequently contribute more births per time unit than those who bear them less frequently. It also provides a future orientation, in keeping with the desire to gauge the potential future fertility of women accepting at a certain age. However, as noted earlier, this choice of mean is subject to an increasing negative bias—mean intervals underestimated—as one moves toward the end of the reproductive period.

"A second approach is based on the reciprocals of marital age-specific birth rates. A basic difficulty here is that some of the women contributing exposure-months to the denominator of the birth rate are already sterile so that an adjustment eliminating the effect of secondary sterility becomes necessary. One way out is to employ the preacceptance birth rates of the acceptors themselves, though here escaping one problem is to court others. If there is a tendency for acceptance to follow closely a birth, then preacceptance birth rates will be biased upwards (their reciprocals correspondingly underestimated) to an extent that varies with the length of preacceptance base period that one selects. One's interest is in potential fertility during a period following acceptance. Therefore one must also adjust preacceptance birth rates to allow for aging and any correlated increase in the likelihood that the women would have been using other methods in the absence of the chosen one. When data allow, one may employ reciprocals of age-specific birth rates of a sample of carefully matched nonacceptors. However, one can question whether an adequate matching of family planning motivation is possible and in any case one is back with the problem that the exposure base of the 'matchees' includes a minority of women already sterile. A further problem, to which Wolfers calls attention, is that for younger ages the inclusion in birth rates of marriage-to-first-birth intervals may be a source of upward bias producing understated durations per potential birth.

"Three problems that exist as serious limitations to the procedures discussed by Wolfers merit further discussion: The acceptor of a program method might have used something else if that program method had not been so available; this likelihood is all the greater if she and her husband had been using contraception earlier. Even if not a previous user, she may have been at the point of starting contraception and therefore would have used another method if the program had not proffered the chosen one. This substitution factor enormously complicates the estimate of potential fertility of acceptors since it means that the estimate of potential fertility should reflect the

natural fertility of the acceptors (including position in the birth interval), as modified by the continuation and effectiveness of the methods (if any) that they might have been using. A mere statement of the problem is enough to indicate its unresolvability in a definitive manner. Thus, any estimate of potential fertility is highly conjectural except in those rare instances when it may be presumed that the alternative to the program method is something close to natural fertility—but even here Leridon would remind us that this is precisely the cultural context in which useful data on fertility are unlikely.

"A second limitation of the procedures reviewed by Wolfers is that while they answer the question of how many births are averted, they do not indicate *when* these births are averted. Effective contraception practiced now obviates births nine months hence; but when discontinued, it may result in a higher frequency of births 9 months later than would otherwise be the case. Wolfers offers a careful definition of net births averted that recognizes the possibility of partially cancelled effects:

> '... births averted is a measure of the quantitative change in the summed expectations of future births to a cohort of women resulting from the adoption or modification of birth control practice by them or their husbands.'

Being calculated is the overall sum of duration-specific deficits and surpluses of births, but not the series of values entering into this sum. More particularly, the procedures are not designed to locate births averted in calendar time, a requirement for any analyst relating acceptor targets to crude birth or growth rates.

"A third limitation is that a satisfactory means of encompassing induced abortion into the births averted calculation has proved elusive. Provision for sterilization presents no formal problem, merely requiring that a zero discontinuation rate be stipulated. But as a practical matter, the long period of 'use' makes more hypothetical any estimate of potential fertility during its practice.

"Wolfers presents in detail three case studies in methodology while alluding to others. One of the three, that of Lee and Isbister, is primarily of historical interest—being an impressive pioneering effort that preceded the availability of life table continuation rates of IUD use and partly as a consequence produced estimates of births averted per device accepted that are now regarded as excessive. A second procedure, by Wolfers himself, conforms as closely as possible to a month-by-month projection of the experience of a cohort of acceptors with various corrections relating to woman-months of effective usage carried along. Reliance is placed on mean prospective birth intervals (births) to measure potential fertility."

[A detailed work-sheet which illustrates how the various corrections and assumptions are incorporated into the monthly calculations is provided in Table 6.2. In the example provided, the proportion of cumulative births averted at the end of three years was 57,516 per 100,000 acceptors for the sub-group considered. Overall, Wolfers' analysis in Singapore found "that one IUD averted (for the first three years after insertion) 0.57 births for Chinese women, 0.65 for Indian and 0.79 for Malay women. Ethnic differences reflected higher fecundity, higher continuation rates and lower previous practice of contraception in the Malay groups. The weighted average was 0.59" (Chapter 6).]

"A third procedure starts with a preliminary estimate of total wearing time, which is then subjected to a series of adjustments in order to derive mean woman-years of effective use per acceptor. This index is divided by an estimate of marriage duration per potential birth to obtain mean births averted per device inserted.

"The last two procedures, though independently developed, are designed to incorporate essentially the same roster of relevant factors. Both procedures were published in 1969. A valuable feature of Chapter 6 is the author's effort to combine the best features of the two methodologies into a single list of recommendations.

"Wolfers lists four 'purposes' which the procedures he examines may serve:

(1) Estimates of births averted provide a necessary demonstration of the way in which program performance leads to demographic change.
(2) Inasmuch as a convenient measure of the economic value of demographic change is the cost (or benefit) of marginal births, program accomplishment measured in terms of numbers of births averted is a useful quantity for cost/benefit analyses of programs.
(3) Births averted estimates expressed as births averted per acceptor of a specific method provide convenient indices for comparison of different program strategies and birth-control methods.
(4) 'Target-setting' or the estimation of the nature and quantity of program input required to achieve pre-set demographic objectives is based on established or estimated relationships between contraceptive use and demographic effects which involve births averted computations.

"The procedures surveyed by Wolfers are perhaps best adapted to the second and third listed purposes. Suitability for the second of

course depends on requisite cost estimates being available. With respect to the third, there are many possible criteria for comparing two contraceptives: safety, effectiveness, continuation, cost, side effects, facilities for dispensing, popularity with younger couples, etc. In the context of family planning evaluation, net births averted per acceptance would appear to be a particularly strategic criterion. It captures in one index several dimensions: continuation, effectiveness, and potential fertility of the acceptors, with the last, if properly measured, reflecting the acceptors' natural fertility as qualified by position in the birth interval at time of acceptance and by the likelihood that alternative contraceptives would have taken the place of the method chosen.

"Yet most comparisons of contraceptives have stopped short at the study of use-effectiveness (i.e., the study of continuation and effectiveness usually by means of a multiple decrement life table analysis). The procedures described by Wolfers have not seen wide use during the half decade since their publication. Part of the explanation may lie in the following dilemma. Only in a relatively few populations such as Singapore, South Korea and Taiwan are the data on fertility and family planning rich enough to permit the methods to be implemented in their full elaboration; but in these populations family planning is widely diffused and couples have many options to program contraception (which itself may include several methods). As a result, any estimates of potential fertility become highly conjectural because of the substitution factor. On the other hand, in situations of incipient family planning when substitution possibilities may be construed as minor, almost certainly one does not have the data for a refined estimate of net births averted.

"Respecting the first listed purpose, there is no question but that if cohort techniques surveyed by Wolfers could be applied to a wide variety of populations, they would teach many lessons about the relative importance of several factors—such as overlap of contraception with secondary sterility or postpartum anovulation, accidental pregnancy, age composition of acceptors, substitution factors and the like—in qualifying the impact of contraception upon marital fertility rates. However, for the reason noted, such a wide ranging study does not seem in the offing.

"A more common approach to investigating the mechanisms of contraceptive impact upon fertility has been through some of the simulation models reviewed by Menken in Chapter 10. Here, complete knowledge is assumed of the parameters pertaining to fertility and family planning. An acceptor cohort (or perhaps a marriage cohort) to which the program method is available is compared with

the same cohort in its absence. Each woman composing the cohort is exposed month by month to relevant probabilities and in this manner followed through her childbearing history. The problem of when births are being averted formally disappears because one is at liberty to compare the current or cumulative *per capita* fertility of the two cohorts at any given duration. Special artifices such as penalties per accidental pregnancy, or multiplication of woman-years of effective use by a measure of potential fertility, become quite unnecessary. In a formal sense, the problem of substitution disappears since the regime of fertility without the program method is being fully specified. Of course there are limitations. One has neither the data nor the knowledge of interrelationships to represent any empirical population faithfully. The models involved in the paired comparisons are always oversimplified representations, with assumptions about parameter values guided in part by data 'borrowed' from other national populations. To a varying degree results are hypothetical because the assumptions underlying them are hypothetical. Moreover, with so many variables contained in the models, it is difficult to draw general conclusions from just a few runs.

"Wolfers' techniques apply to cohorts of acceptors, not period populations, and are therefore not directly designed for the fourth listed purpose, namely, assessing the annual age schedules of acceptance of a specified mix of program methods necessary to achieve predesignated reductions in the crude birth or crude growth rate. A simulation model such as POPSIM that represents a period population and includes a family planning module might be utilized; but as a practical matter, the expense and data requirements of such an elaborate model have so far debarred its wide use for this purpose. As Ross in Chapter 3 documents, the leading approach has been component projection designed so that annual marital age-specific birth rates come to reflect assumptions about acceptance, continuation and effectiveness of contraception and about the relative potential fertility of acceptors versus married nonacceptors.

"Such a component projection may be viewed as the imbrication of annual cohorts of acceptors, classified by age, together with an age-graded pool of eligible nonacceptors. The problem of when births are averted is usually handled by multiplying woman-years of effective use during the previous calendar year, in each age class, by an appropriate age-specific potential fertility rate, thereby deriving births averted this year in each age class. Averted births are then subtracted from the potential births to be expected in the absence of the program method.

"Seen in this perspective, many of the issues addressed in

Chapter 4 retain full pertinence. For example, age-specific potential fertility rates should reflect substitution effects and the lower prevalence of secondary sterility among acceptors. Annual woman-years of contraceptive use should be refined to exclude overlap with anovulation or sterility and to make allowance for accidental pregnancies.

"In none of the schemes for assessing numbers of acceptors needed to achieve predesignated demographic goals that are reviewed in Chapter 3 are all the refinements proposed in Chapter 6 incorporated. Yet the reasons for omission are rarely disavowal of relevance, but such reasons as lack of data, inadequate computational facilities, or because of the algebraic constraints stemming from the greater complexity of a period population as compared to a single acceptor cohort. A basic value, then, of the rather elaborate procedures discussed by Wolfers is providing a set of procedural standards to be emulated in applied target-setting schemes as data and facilities permit."

VI. CONTROLLED TRIALS, MATCHING STUDIES, AREAL ANALYSIS AND SURVEYS

The foregoing review of births averted analysis highlights two particularly difficult problems which arise in pursuing this approach. One is the substitution problem which includes the possible replacement by program methods of birth control methods used outside the program, as well as the possibility that those who become acceptors from the program would have adopted private means in the absence of a program. The second problem is the difficulty of ascertaining the potential fertility of acceptors, particularly when social and economic conditions are changing rapidly. At the same time, births averted analysis makes no allowance for a program's indirect effect in stimulating the target population to adopt birth control means through private channels, though theoretically this effect could be incorporated if its magnitude were known.

For these, as well as other reasons, other attempts to measure the impact of a program on fertility have been developed. The hallmark of these approaches is the focus on observed fertility levels of the population or select groups within it and the determination of the relative influence of the program vis-à-vis other factors affecting fertility. This determination of program effect may be sought by controlling for the non-program factors or by comparing the relative importance of program effects with those from other sources. Chapters 7, 8 and 9 of this volume describe a variety of techniques illus-

trative of this broad approach. These techniques, based on experimental design and matching studies, and multivariate analyses of areal or survey data, are part of the repertoire of scientific investigation and not unique to demographic analysis or program evaluation. As such, and given their detailed treatment in the respective chapters, the discussion here will be shorter than the previous sections.

EXPERIMENTAL DESIGN AND MATCHING STUDIES

Theoretically, controlled experimental trials would be the technique *par excellence* for measuring the demographic impact of a family planning program. Practical, ethical and political considerations, however, invoke serious restraints on their use for this type of program. On an individual level, random assignment of women to accepting and non-accepting groups makes little sense. Experimental design with areas as the units of interest is more feasible but has been employed to only a limited degree in studying the effects of a program on fertility. The Khanna (Wyon and Gordon, 1962) and Singur (The Population Council, 1963) studies represent examples along these lines. As Wells' discussion in Chapter 7 notes, it may be unethical to withold services considered beneficial from any part of the population; in addition, there are political considerations which lead policymakers to initiate a program throughout a country; and from a practical standpoint, it is often difficult to control communication between experimental and control areas, even if the previous two restraints were overcome. Additional difficulties noted arise from the need for replication to establish the degree of experimental error, the problem of selecting areas so that they are representative of the population, the possibility of confounding from the actions of the personnel employed in the experimental areas, and the problem of reactive measurements in monitoring the progress of the experiment.

Even though "true" experiments for the purpose of measuring demographic impact are difficult to achieve, they serve usefully as ideal-types against which other designs can be evaluated for possible biases. In addition, they can be successfully employed in other phases of family planning programs. For example, experiments have been used in the city of Taichung to study the cost effectiveness of different types of program efforts (Freedman and Takeshita, 1969); in studying the effect of different types of IUDs (Tietze and Lewit, 1972); and in investigating the returns from an incentive payment system to field workers (Chang *et al.*, 1972). One tends to agree with Wells' judgement that controlled experiments are not being utilized

to the extent possible in the evaluation of various phases of family planning programs.

Most analyses of program impact must rely on observational studies, where units have not been randomly assigned to treatments. In such situations, one attempts to demonstrate that disturbing variables, other than the treatment, do not account for the observed results. In the framework of Chapter 7, one enumerates and classifies the possible disturbing variables, and on this basis removes or adjusts their possible biasing effects by matching (pair-wise or mean matching), adjustments of sub-group means, or covariance adjustment by regression. The relative merits of each approach depend on the nature of the disturbing variables.

Chapter 7 presents a detailed analysis of the method of accounting for disturbing variables through matching. There are three important issues:

(1) the method of forming the comparison group,
(2) the disturbing variables considered, and
(3) the period to be covered.

We illustrate these with reference to a matching study which seeks to compare the pre- and post-acceptance fertility of acceptors with a suitable control group of nonacceptors. The first issue is that the formation of the two groups must be specific to a point in time, and in particular, that later acceptance by those starting out in the control group is not grounds for removing them from the comparison. Perhaps the most serious problem facing matching studies is in the choice of possible disturbing variables to serve as the criteria for matching. In the absence of an adequate model of fertility behavior, one can never be certain that some omitted disturbing variable might not account for the observed difference. An example of a rather subtle disturbing variable was described in the previous section. If there is no matching for open birth interval, or improper matching, the control group may have a higher frequency of secondary sterility than the acceptor group. A more obvious omission from this type of study concerns the motivation of the acceptors and non-acceptors. If the purpose of the study is to show that acceptance makes a difference independently of socio-economic trends, one must control for the possibility that acceptors are simply more sensitive to these trends. This might be termed a "selection-history" interaction and serves as a competing hypothesis to the effect of acceptance (Campbell and Stanley, 1963: 48).

Another competing hypothesis arises from the difference between the fertility levels of acceptors and that of the population. Insofar as acceptors, at the time of acceptance, generally have higher fertility than the general population, then "regression-effects" can occur, in that both acceptors and those selected as matches will be "extreme" members of the population and likely to show lower future fertility due to chance factors. (Campbell and Stanley, 1963: 49). Note that in this case, the "regression-effect" operates on both groups and may not contribute any "bias" to the comparison and that the failure to control for secondary sterility may lead to smaller differences in the post-acceptance fertility comparison, and thus underestimate the impact of acceptance. Other omitted variables may, of course, produce an opposite effect.

The third issue is the need to decide on a proper time perspective for comparing the post-acceptance fertility of acceptors and controls. Short term differences may disappear over a longer period, as illustrated in Chapter 7. This problem arises with experimental designs as well as matching studies. Both the Khanna and Singur studies showed that differences between experimental and control villages changed considerably over time.

Some of the difficulties encountered in matching studies, like the need to fully account for disturbing variables, are relevant to other ways of analyzing observational data. Some, like the regression effect, are associated with the matching technique and introduce additional possibilities of bias. Thus the results of a single matching study must be interpreted with caution. However, the potential of matching studies for the study of demographic impact can be improved through some of the suggestions described in Chapter 7—increasing the number of comparison groups, adding potential disturbing variables, and collecting longitudinal data. Also, given the weakness of any observational approach, a wise strategy would be to employ quite different designs to see if they result in similar inferences.

MULTIVARIATE ANALYSIS OF AREAL DATA

One alternative to matching designs is the application of regression analysis to either areal or individual data. Chapter 8 discusses the analysis of areal data and the following review of the advantages and disadvantages of this approach has been prepared by H. B. Wells.

"*Main uses and range of applicability.* The main uses of multivariate areal data are the same as for other analyses: to seek evidence that program inputs are positively correlated with and indeed

re causally related to achievement of program goals. In the absence f controlled trials using areas as experimental units, results from a ingle multivariate analysis may be utilized in combination with other tudy results to develop plausible models of causal relationships.

"A geographic area is the appropriate unit for observation and nalysis when interest focuses on the total aggregate of individuals ssociated with the area rather than individuals within the area. For xample, one can hypothesize that, 'other things being equal', areas ith high family planning program input would show more fertility eduction than areas with low program input.

"If suitable data on program input and other measures presumed b be related to fertility are jointly available for the same set of areas, number of multivariate techniques of analysis, including regression nalysis, may be applied to estimate relationships between indepen- ent variables, including program input, and one or more dependent ariables, including fertility measure(s). Univariate multiple regres- ion allows one to estimate the impact of program inputs on fertility, olding 'other factors equal'. Path analysis provides a means of urther analyzing and representing relationships implicit or explicit in ultiple regression analysis. Factor analysis is another technique or analyzing relationships between sets of variables and other sets f variables rather than between a single variable and other (sets of) ariables. Multivariate multiple regression, in which multiple depend- nt variables are jointly regressed upon sets of the same independ- nt variables, is another technique which can be applied to such ata. Chapter 8 restricts itself primarily to univariate multiple regres- ion and its relationships to path analysis, although some examples re given of the application of other methods.

"Regression models can either be applied separately to data for ndividuals or to data for areal aggregates, or both individual and real characteristics may be included in the same model, although the ombined model is mentioned here only in a reference.

"The range of application of multivariate techniques to areal ata is limited to situations where data of adequate quality are vailable for all variables specified as required for the model at hand.

"*Conceptual basis.* Regression analysis using areal units for bservation must be based upon a conceptual model of how depend- nt variables for an area, such as pregnancy rate, fertility rates, or cceptance rates, are related to independent variables for the area, ncluding endogenous measures of family planning program input, nd exogenous non-program variables for the area. The conceptual odel is then translated into a statistical equation which can be pplied to the data to estimate relationships.

"The aim of analysis will be to estimate the relative impact, i.e. the causal effect of the program on the fertility measure, after removing or adjusting for the influence of non-program variables.

"Relationships which hold at the areal level may suggest the need to investigate relationships based upon individuals but cannot be considered to hold at the individual level until verified.

"*Data needs and accessibility.* Useful non-program data may be obtained from census and civil registration sources when these are available and, as such, are basically cross-sections or repeated cross-sections by nature. Program input data can be obtained as a by-product of program operation, provided that adequate planning is done to assure that area units for compilation are the same as, or can be combined to correspond with, areal units used in other data systems. In the absence of, or as a supplement to, adequate census or civil registration data, survey data collected from a sample of areal units may also be used in multiple regression analysis. KAP measures of knowledge, attitudes and practice—especially practice outside the program, are potentially valuable.

"In view of the large number of variables theoretically related to fertility, it appears unlikely that areal level data for all of them will be available for realistic models in the immediate future, although with sufficient planning they can become available eventually.

"*Major advantages and limitations.* The major advantage of the multiple regression model is that it is conceptually capable of indicating what variables are associated either with changes in fertility or with a program response variable such as acceptance rates.

"Disadvantages are many. It is usually necessary to restrict the model to overly simple representations involving a limited number of macro-level independent variables and the dependent response variable at the areal level. Further, as indicated by Hermalin, the data which are available often violate the assumptions of the statistical model. When this occurs, statistical tests and estimates of relationships at best are biased to an unknown degree and at worst may be misleading.

"*Interpretation and use of results.* In view of the potential pitfalls due to the types and quality of areal data either available or collectable, caution should be the watch-word in using the results of regression analyses. One should use such results as provisional and indicative of possible causal relationships. To the extent that analysis of new or alternative data sets, taking adequate account of time lags and possible response errors, confirms the same or related causal effects, the evidence from a single analysis takes on added weight.

"It is clear then, that single unconfirmed study results are not

recise and should be considered as an hypothesis to be tested rther, and hence should be translated into administrative decisions nly after careful consideration of alternative explanations.

"*Chief lines of further work*

(a) Technical problems: A major unresolved problem is the evelopment of methods of estimating the possible extent of bias resent in results of multivariate analysis when the necessary assumptons are violated. This might indicate conditions under which results ould be 'acceptable' for decision making.

"A related problem is the development of methods of data collecon and processing which either eliminate measurement errors or ermit measurement of such errors and methods of adjustment for em.

"More research is required to determine whether weighting areal bservations by some measure, for example, the number of females resent, would result in better estimates.

"There is a need to develop models which take into account patial as well as temporal correlations in the variables. In this onnection, more investigation of the pooled cross-section approach indicated.

"As demonstrated by Hermalin, different analysts using the same ata sets can obtain different results. More experiments of this type re desirable in order to begin to quantify the magnitude of analyst ariation on results of social experimentation. This may be especially nportant in multi-country collaborative studies.

"(b) Cross-cultural comparisons: It is desirable to establish ulti-country studies in which the same conceptual models, definitions nd procedures are applied to different data sets. If, in addition, nalyst variation studies as described under (a) above are underaken, it may be possible to see if a single causal model holds across any different family planning programs.

"(c) Validation from other analyses: Covariance adjustment of bservational data is another variation of the multiple regression pproach. The other techniques of analysis of observational data escribed in Chapter 7 should yield similar results when applied to e same data set. This confirmation would in no way validate the esults of regression analysis. Validation of relationships can only e done through replication with other data sets and by testing a ide range of alternative theories of causal relationships."

THE USE OF RETROSPECTIVE SURVEYS

The sample survey covering fertility and contraceptive practices is probably the most versatile technique for studying various phases of a family planning program. Its use as an alternative method of generating data on acceptances and for validating service statistics has already been noted. Recognition has also been taken of its importance as a source of data needed by other techniques of evaluation, and as a possible frame from which subjects for various types of follow-up or matching studies can be drawn. Chapter 9 points out a number of additional uses:

(1) Measuring current fertility levels and changes in fertility. These may not be available from other sources.
(2) Obtaining attitudes and intentions with regard to future fertility.
(3) Description of the population with regard to other basic demographic variables and in terms of socio-economic characteristics.
(4) Providing baseline data on knowledge, attitudes and practice of contraception, and changes in these variables over time.
(5) Establishing degree of contact with various aspects of the program—clinics, fieldworkers, etc.—and reactions to these contacts.
(6) Measuring the proportion of contraceptive use coming from the private sector.

Some of these measurements can be obtained from a single KAP-type survey, while others require a series of surveys. These data provide program administrators with a number of valuable insights:

(1) Surveys taken before or at the start of a program indicate potential levels of receptivity to the program in total and help identify groups within the target population who may be particularly interested in the services provided. At the same time they provide baseline data on knowledge, attitudes and practices against which to compare future distributions, and thus help monitor the progress of the program.
(2) Data on utilization of and satisfaction with services often point out needed changes in program operation and help distinguish the more successful from the less successful program strategies.
(3) Analysis of the data indicates the extent of socio-economic differentials on fertility and contraceptive behavior and helps identify the relationships requiring further study.

4) Interrelationships of past and current use of contraception, along with sources of the contraceptives and contact with the program, can provide clues as to the magnitude of the substitution effect and of the indirect effect of the program.

When KAP surveys are repeated at periodic levels, their value ıcreases since one obtains not only a picture of the changing atterns but can begin to study the factors associated with these ıanges. An example of the insights generated by successive surveys is provided by Freedman *et al.* (1974). Successive surveys can e based on independent samples or can involve reinterviews with ıe respondents of an earlier survey. The latter approach, referred o as panel studies, often permits of more detailed analysis of changes ı attitudes and behavior. For example, those who do or do not ıange their attitudes or practices concerning contraception can be lentified and the factors associated with such changes studied. espondents' intentions with regard to future fertility at the first urvey can be compared to their subsequent behavior to see to what xtent intentions plus other characteristics are predictive of future ehavior. Over time, some surveys can be of the panel type while thers are based on independent cross-sections.

When one focuses more narrowly on assessing the impact of a rogram on fertility, the value of a survey arises from its ability o provide data on a large number of variables that affect the process. he general methodology is similar to that described for areal multiariate analysis. That is, one requires a measure of fertility, which erves as the dependent variable, measures of socio-economic and ther characteristics which are thought to be determinants of fertility, nd measures of acceptance or contact with the program to serve s program influences. The choice of variables is determined by a onceptual model of fertility behavior. In addition, one must use ppropriate statistical techniques to estimate the importance of the rogram variables vis-à-vis other fertility determinants. Possible conceptual models and problems of statistical estimation are reviewed in hapter 9. As compared to areal analysis, multivariate models on the ıdividual level have an advantage in that the analyst, through the urvey, can collect data on the variables deemed relevant and does ot have to rely on the limited characteristics usually available by rea. In this way, more elaborate models can be tested. Counteralancing this advantage to some extent is that survey data are more ubject to certain types of errors—in terms of sampling and response rrors—than areal data. If surveys are to be used to test the demoraphic impact of a program, it is important to develop the conceptual

model prior to the execution of the survey. Any given survey is a cross-section at one point in time, and if one is to order a set of variables in causal or temporal sequence, it is often necessary to elicit the proper ordering of various attitudes and behaviors through the proper design of the questionnaire.

Multivariate analysis of data from surveys is an important technique for utilizing observational, as distinct from experimental, information. As was the case with the other observational techniques reviewed in this section, a single test of one model is apt to be less revealing than tests repeated over time and place with alternative models.

VII. OTHER APPROACHES

Chapters 11 and 12 illustrate some other approaches to determining the demographic impact of a program which involve studying the trend of fertility before and after the start of a program. The Singapore case study, for example, traces the percentage decrease in the total fertility rate over the period 1957-1970, showing that percentage declines were substantially greater after the program started (Chapter 11, Table 11.19). D. Wolfers (discussion paper) comments that basing the change on the previous year's total fertility rate may underestimate the impact of the program. His comments follow:

> "The conventional methods of display tend always to short-change a program. Thus Table 11.15 shows the acceleration in the decline in fertility rates which accompanied the innovation of the program. The method used is to express each year's total fertility rate as a proportion of the previous year's and assess the percentage decline by year. This method shows a large decline in 1966-67, as would be expected with a successful program, but a speedy return in the next three years to the moderate rate of decline prevailing before the introduction of the program. In fact, this rapid return to a low rate of change is a statistical artifact deriving from adherence to the convention of a zero base line. Just as the weight loss of a starving man will become both proportionately and in absolute terms less each day as his weight approaches an irreducible minimum, which is not zero, so will birth rates, death rates and other measurements which do not have the potential to vanish altogether. The zero level for fertility measurements ought to be the target level for the program (D. Wolfers, 1970). In the case of Singapore this is implicitly replacement reproduction, corresponding at current mortality rates to a total fertility rate of about 2,200. The application of this base line would allow us to replace Table 11.19 with Table 13.2, showing the true phenomenon of a very rapid decline fully sustained over the first three years of the program, with 61 per cent of the excess fertility present in 1966 being shed by 1970."

Table 13.2. *Excess total fertility over replacement, 1957-1970*

Year	*Excess total fertility (a)*	*Annual percentage decline in excess fertility*
1957	4,265	—
1958	4,222	1.0
1959	4,030	4.5
1960	3,668	9.0
1961	3,306	9.9
1962	3,058	7.5
1963	2,999	2.0
1964	2,794	6.8
1965	2,586	7.4
1966	2,306	10.8
1967	1,765	23.5
1968	1,374	22.2
1969	1,057	23.1
1970	904	14.5

(a) Actual total fertility minus replacement fertility of 2,200.

For Taiwan a somewhat more elaborate trend analysis has been employed in which trend lines have been fitted by least squares procedures for the periods before and after the start of the program (Chapter 12, Figure 12.4 and Table 12.25).

Both types of trend analysis might be said to provide presumptive evidence for the impact of the program, but not conclusive demonstration. In neither approach are the social and economic trends taken into account explicitly. Rather their impact is assumed to be captured by the pattern of fertility change prior to the program's inception. Another difficulty with trend analysis is that the imputed magnitude of effect is often sensitive to the time points chosen.

Another type of trend analysis is possible through data collected from a survey. If the age span of the adult respondents in a survey is expanded beyond the reproductive ages, one can measure the changes over birth and marriage cohorts in the proportions using contraception and in the methods used. The time series thus derived can be compared with measures of fertility, suitably lagged, obtained either from the survey or from vital statistics. This approach introduces the level of contraception explicitly into the analysis and provides a measure of association between it and fertility. If this relationship is strongly negative, it strengthens the presumption of

a program effect to the extent that the program can claim credit for increasing the levels of contraceptive use. The Lewis-Faning study mentioned in Chapter 9, among other objectives, derived time series of contraceptive use in England.

In both the Singapore and Taiwan case studies, there is a decomposition of the change in a fertility measure (e.g., crude birth rate) attributable to various demographic components (e.g., age distribution, nuptiality, marital fertility). Such a practice is an important safeguard in any kind of trend analysis in that it prevents falsely attributing to the program that which is due to change in demographic structure. In addition, such analysis helps identify the maximum amount of change that could have arisen from the program.

Another type of demographic analysis, presented in Chapter 11, is that of the gross total fertility rate by birth order. This measure is the sum of the age-specific fertility rates for births of a given order computed by five-year age interval. From 1967 to 1970, this rate shows small declines in lower order births but steep declines in higher order births; in addition, "major reductions have increasingly shifted orders of birth" (Table 11.20). The authors interpret this pattern to reflect the role of childspacing, "whereby early reductions in fertility... are somehow 'compensated' for later as the delayed births appear." There is some question, however, whether the gross total fertility rate is fully adequate to capture this effect. As Wolfers (discussion paper) notes:

> "While the denominators for the gross total fertility rates are all women of the specific age-group in the population, the women at risk are only those of parity less than the birth order. If therefore in any given year there is a deficiency of births of order n, then in the next-but-one year (assuming few birth intervals of less than one year) and subsequent years, there will be a deficiency of mothers of parity n and hence an inevitable deficiency of births of order $n + 1$. At the same time there will be an excess of mothers of parity $n - 1$ which will soon lead to a compensatory increase in births of order n. This latter effect is of course the postponement effect to which the authors refer. In some ways, therefore, this form of analysis is more difficult to interpret than parity-progression ratios, in which the denominator of the n-order specific rate is women of parity $n - 1$."

The development of family planning programs around the world has opened up the possibility of international comparisons as a way of assessing their impact. In one such attempt, Berelson (1974) cross-classifies countries by strength of program effort and by institutional setting, based on a number of socio-economic factors deemed relevant to fertility, and then examines the acceptance rates and the changes in birth rates in the countries in each cell. He finds that

both the strength of program and the institutional setting affect fertility decline and acceptance level. In another comparison, Ravenholt and Chao (1974) trace the pattern of change in age-specific fertility rates over time in a group of countries with "vigorous" family planning programs and a select group of countries without such programs. The countries in the former group generally show substantial declines in the 1960-1970 decade, while those in the latter group show little or no decline. In addition the authors show that some of the countries without "vigorous" programs and with little change in fertility, had economic growth equal to or greater than those in the other group.

VIII. CONCLUSIONS

An impressive array of sophisticated techniques has been developed or adapted for the purpose of providing insights into the demographic impact of a family planning program. Though each makes some contribution, none answers the complex problem fully. Though each possesses some advantages not shared by others, none is free of limitations.

The types of limitations may be subsumed under a few broad headings. One is the demands made in terms of data. Some methods, such as births-averted analysis, require measurements on a number of aspects of the reproductive process—such as secondary sterility, post-partum amenorrhea, etc.—that are not often available. A comparison of the factors called for in Chapter 6 with the inputs used for target setting in Chapter 3 point out these differences. Areal analysis would benefit from a richer variety of characteristics per area to permit more elaborate model-building, but for many countries, even a basic set of variables does not exist. Another broad problem area, somewhat distinct from data availability, concerns the ability to operationalize key concepts in a given approach. Thus, in births-averted analysis, the data required to capture the substitution effect and the potential fertility of acceptors are not easy to specify. Another need is to develop knowledge of the range of precision that can be expected from observational studies. As indicated, such studies do not always meet the statistical assumptions on which the technique rests and it is therefore important to have some idea of the magnitude of the possible biases.

From a somewhat different standpoint, many of the limitations can be viewed as arising from the dimension of time. Thus, life table analysis of contraceptive use is limited by the need to extrapolate

from short run observations to long run patterns of use. Similarly, the problem of potential fertility can be viewed as a problem of extrapolating the past fertility of acceptors into some future period. In areal analysis, and perhaps to a more limited extent in survey analysis, one faces the problem of converting cross-sectional observations into a dynamic model. In births-averted analysis, one would like to convert cohort estimates into period measures. On a more general level, many of the techniques of evaluation are faced with the twin problems of estimating long range impact from short range studies, and of detecting short range impact in long range programs.

Improvements in the power of the techniques will come from attention to the limitations described in this volume, and from cumulative experience in their application. As to the latter, there appears to be a need for more work with the same technique in different countries. In this way the sensitivity of the results to differences in the magnitude of the various components can be developed. In addition, as Srikantan (discussion paper) notes, cross-country studies will help establish the interrelation between socio-economic change and effectiveness of the family planning program. It is possible that a program is more effective in certain socio-economic and demographic settings than in others. Many of the techniques described have been sparsely used and in few cases is there the opportunity to compare the results of the same technique in different settings.

Also needed is the application of a number of different techniques within the same country, best exemplified by the work in Taiwan. Given the limitations of any one technique, stronger inference about the demographic impact of a program will be possible when distinct approaches reach the same conclusion. Techniques of evaluation will also benefit from basic research advances in our understanding of the determinants of fertility levels and change.

Little in this volume speaks directly of the extent to which research results are utilized in program administration. It is appropriate to conclude, however, with a few observations about how the researcher and policy maker can facilitate each other's work. From the administrator's viewpoint, one may ask if there is a proper match between the questions posed and the research results. Akin to this, are the research results capable of translation into factors which can be manipulated to the advantage of the program? Does the research technique demand data of a quantity and quality that cannot be reasonably obtained in the current setting?

From the researcher's viewpoint, some important questions revolve about the specification of the criteria for success or failure of a program. The policy makers need to specify the magnitude as

well as the nature of the impact they are seeking, and the time perspective within which the effect is to be realized. Is excessive refinement being called for, which is beyond the range of current techniques and data systems? Or is it possible that researchers are overlooking useful but approximate techniques in the mistaken notion that they are insufficient for administrative purposes? Better coordination between policy makers and researchers should produce more useful results to the former and guidance to the latter for further development of the tools of evaluation.

REFERENCES

Austin, James E.
1973 "The management bottleneck in family planning programs", Studies in Family Planning 4 (12): 343-350.

Berelson, Bernard
1974 "An evaluation of the effects of population control programs", Studies in Family Planning 5 (1): 2-12.

Campbell, Donald T. and Julian C. Stanley
1963 Experimental and Quasi-Experimental Design for Research, Chicago, Rand McNally and Company.

Chang, M. C., George P. Cernada and T. H. Sun
1972 "A field-worker incentive experimental study", Studies in Family Planning 3 (11): 270-272.

Freedman, Ronald and John Y. Takeshita
1969 Family Planning in Taiwan: An Experiment in Social Change, Princeton, Princeton University Press.

Freedman, Ronald, L. C. Coombs, Ming-Cheng Chang and Te-Hsiung Sun
1974 "Trends in fertility, family size preferences, and practice of family planning: Taiwan, 1965-1973", Studies in Family Planning 5 (9): 270-288.

Lapham, R. and W. Parker Mauldin
1972 "National family planning programs: review and evaluation", Studies in Family Planning 3 (3): 29-52.

Liu, P. T., L. P. Chow and H. Abbey
1972 "A study of IUD retention by curve-fitting", Demography 9 (February): 1-11.

Mauldin, W. P., D. Nortman and F. Stephan
1967 "Retention of IUDs: an international comparison", Studies in Family Planning (18): 1-12.

McCorquodale, Donald W. and Thomas W. Pullum
1974 "A mathematical model for determining effectiveness of family planning clinics", Studies in Family Planning 5 (7): 232-237.

Population Council
1963 "India: the Singur study", Studies in Family Planning, (1): 1-4.

Ravenholt, R. T. and John Chao
1974 World Fertility Trends, 1974. Population Report, Family Planning Programs, series J (2).

Reynolds, Jack
1972 "Evaluation of family planning program performance: a critical review", Demography 9: 69-86.

Suchman, Edward A.
1967 Evaluative Research. New York, Russell Sage Foundation.

Tietze, C., and S. Lewit
1972 "Comparison of the copper T and loop D: a research report". Studies in Family Planning 3 (11): 277-278.

Wishik, Samuel and K. H. Chen
1973 "Couple-year of protection: a measure of family planning program output" (Manual No. 7 of a series), Manuals for Evaluation of Family Planning and Population Programs. Columbia University, International Institute for the Study of Human Reproduction.

Wolfers, D.
1970 "The Singapore family planning program: further evaluation data", American Journal of Public Health 60: 2354-2360.

Wyon, John B. and John E. Gordon
1962 "A long term prospective-type field study of population dynamics in the Punjab, India", in Clyde V. Kiser (ed.), Research in Family Planning. Princeton, Princeton University Press.

CONTRIBUTORS

ROGER C. AVERY
International Population Program
323 Uris Hall
Cornell University
Ithaca, New York 14850

C. CHANDRASEKARAN *
Senior Advisor on Policy and Planning
United Nations Fund for Population Activities
485 Lexington Avenue
New York, New York 10017

C. T. CHANG
Lecturer
Department of Sociology
University of Singapore
Bukit Timah Road
Singapore 10
Singapore

PHILIP M. HAUSER
Director
Population Research Center
University of Chicago
1126 East 59th Street
Chicago, Illinois 60637

ALBERT I. HERMALIN
Associate Director
Population Studies Center
University of Michigan
1225 South University Avenue
Ann Arbor, Michigan 48104

GWENDOLYN JOHNSON-ACSÁDI
Chief, Fertility Studies Section
Population Division
United Nations
New York, New York 10017

HENRI LERIDON
Institut National d'Etudes Démographiques
27, rue du Commandeur
Paris 14
France

W. PARKER MAULDIN
Acting President
The Population Council
245 Park Avenue
New York, New York 10017

JANE MENKEN
Office of Population Research
Princeton University
5 Ivy Lane
Princeton, New Jersey 08540

ROBERT G. POTTER
Professor, Department of Sociology
Brown University
Providence, Rhode Island 02912

JOHN A. ROSS
Associate Director
Technical Assistance Division
The Population Council
245 Park Avenue
New York, New York 10017

K. SRINIVASAN
Population Officer
Population Division
United Nations Economic and Social Commission for Asia and the Pacific
Sala Santitham
Bangkok, Thailand

T. H. SUN
Executive Secretary
Committee on Family Planning
Taiwan Provincial Health Department
Taichung, Taiwan

H. BRADLEY WELLS
Professor, Department of Biostatistics
School of Public Health
University of North Carolina
Chapel Hill, North Carolina 27515

DAVID WOLFERS
International Institute for the Study of Human Reproduction
Columbia University
78 Haven Avenue
New York, New York 10032

* Currently affiliated as Population Specialist with the International Bank for Reconstruction and Development, Djakarta, Indonesia.

INDEX

Abbey, R., 451, 453, 503, 553
Abernathy, J. R., 111, 128
Abortion, 116, 152, 159, 167, 172, 184, 207, 319, 355, 361, 365; as method of birth control, 3-4, 18, 95, 226, 308, 316, 336, 368, 509; births averted through, 80, 164, 191-192, 535; disguised as miscarriage, 110, 117; in specific countries, 407-408, 77, 308, 455, 460, 470-472, 475, 487, 495-496; incidence of, 110-111, 337-338, 359; legalized, 2, 66, 110, 362, 397, 407; sterility resulting from, 117, 124, 174
Abstinence, 172, 336
Acceptance, 369, 372, 550-551; definitions of, 17-21, 26, 44-46, 507-509; simulation models of, 177, 358, 362
Acceptance rates; cumulative, 68, 358, 512; by demographic characteristics, 22-23, 28-34, 37-43, 48-49, 83-84, 142-143, 156, 177-178, 190, 239-240, 333, 363-366, 511; initial, 221; methods of estimating, 231-235; by socio-economic characteristics, 22, 34-38, 43, 48, 190, 206, 331-332, 363, 399-400; time factor in, 83, 233-234, 364
Acceptors, 12, 27-29, 370, 511; age of, 22-23, 57, 61, 68, 171, 176, 180, 190-192, 368-369, 374-376; age-specific fertility rates of, 177-178, 190, 231, 278; attitudes of, 4-7, 11-13, 18, 264, 302-303, 307-310, 319-320, 328, 331, 472-473; characteristics, 12, 27-29, 370, 511; education of, 23-24, 30, 34-38, 43, 48, 180, 308; ethnic background, 36-37, 43, 196-203, 206; fecundability, 10-11, 174-177, 358; in specific countries, 30-31, 40, 191, 393-394, 399-408, 420-442, 74-77, 85, 431, 436-437, 451, 455-468, 474-475, 480-483, 488-489, 495-500; marital status, 28-29, 48-49, 64-67, 84, 234, 366; motivation, 11, 18, 142, 242, 316, 399, 408, 450, 472, 495, 541; number of living children, 22-23, 29-43, 57, 83, 142-143, 167, 172, 177-180, 207, 361-368, 399-400; occupation, 23, 48, 226, 26-62, 289, 307, 332, 339, 399, 432-433; potential fertility of, 10-11, 22, 57-58, 61, 68, 83, 164, 172, 176-177, 190-193, 208, 233, 529-539, 551-552; pre-program contraceptive history, 12, 34, 57-58, 169-172, 190, 209, 226, 231, 265, 399, 510; withdrawal from program, 61, 64, 71, 84, 139, 156, 168, 204-205, 233-234, 240-241, 375
Acceptor data, 21-25, 34, 42-46, 50-51, 134, 169, 179-180, 434-435, 443, 492, 506-511
Acceptor targets, 12, 26-27, 48, 51, 55-58, 64-65, 366, 394-398, 508-509, 526-528, 539
Adil, Enver, 89
Adlakha, A., 52, 179, 211
Agarwala, S. N., 186, 192, 211, 249, 294
Alan, J., 337, 348
Algeria, 117
Alker, Howard R. Jr., 248, 294
Analysis of variance, 330-331
Anderson, David L., 275, 298
Andrews, G., 331, 347
Ann, T. B., 244
Areal analysis, 289, 346-394, 501, 552; operational problems of, 256-257, 292-294, 544-545; survey data and, 309, 330, 547-548
Attitudes toward contraceptive practice, *see* Contraceptive practice
Austin, James E., 507, 553
Autocorrelation, spatial, 275-277
Avery, R. C., 139, 153, 160, 206, 211, 450-453, 457, 501, 516-517, 520-523

Balakrishnan, T. R., 150, 160
Bali, 62

Bang, S., 349
Bangkok, 22-23
Bangladesh, 335-336
Barrett, J. C., 128, 174-175, 186, 205-211, 354, 371, 375-376
Baum, S., 306, 347
Bean, Lee L., 7, 13, 70, 89
Berelson, B., 3-7, 13-14, 88-89, 221-222, 243, 302, 347, 431, 434, 501, 550, 553
Berent, J., 5, 14
Berry, Brian J. L., 277, 295
Beta coefficients, 282
Bhatti, A. D., 70, 89
Bierman, J. E., 112-116, 128, 131
Billewicz, W. Z., 227-229, 243
Birth certificates, 234-236
Birth control, practice of: *see* Abortion and Contraceptive methods and use
Birth intervals, 62, 191; age-specific distribution of, 98, 122, 180; family planning and, 34, 38, 43, 83, 150, 174, 318, 374; fertility rates estimated from, 96-98, 122-123, 171-173, 178-181, 208, 239, 315-317, 529-534; length of, 22, 28, 120-123, 167, 177, 203-205; nonsusceptible period and, 102
Birth order, 57-58, 97, 120, 177-178, 204, 415-416, 432, 550
Birth rate: age-specific, 268, 366; decline in, 51, 86, 362, 375, 383-388, 550-551; estimation of, 57, 328, 369; fecundability and, 164, 363-364, 529-532; fluctuations in, 353-355, 358-359, 370-371; in specific countries, 59, 61, 67, 70, 78-80, 84, 268, 383-385, 388, 394-395, 398, 412-413, 419, 430, 476-478, 483-489, 495-499; marital fertility and, 280, 366, 410-413; simulation models of, 352, 358-359, 366; variables affecting, 202, 262-263, 268-274.
Birth registration, 123, 302, 432-433
Birth statistics, 264
Births averted, 12-13, 49-51, 88, 507, 551; abortion and, 164, 191-192, 535; contraception and, 88, 134, 164-166, 174, 193-194, 195-206, 210, 313, 324, 510, 522-539; family planning programs and, 163-164, 217, 225, 230; fecundability and, 164, 529-532; in specific countries, 62-64, 68, 73, 79-87, 95-96, 192, 196-203, 206, 417, 459, 465-469, 481-482, 495-496, 499, 536
Births postponed, 62, 174, 368
Bi-variate analysis, 329-330
Blake, Judith, 311-312, 347, 351
Blalock, H. M. Jr., 260, 288, 295
Bobo (Upper Volta), 118
Bodmer, W. F., 363, 375-376
Bogue, D. J., 5, 14
Bombay, 33, 47
Boynton, John W., 245, 296
Brass, William, 177, 207, 211
Breastfeeding, *see also* Lactation: contraceptive function of, 119, 336; customs of, 128; duration of, 88, 119-120

Campbell, Donald T., 541-542, 553
Cantrelle, P., 131
Caribbean, 179
Caro, F. G., 7, 14, 217, 243
Catholicism, fertility behavior and, 305
Cazenave, M., 121, 131
Celibacy, 5, 95, 311
Census data, 29-30, 65, 179, 263-266, 289, 301, 338, 381, 385, 392, 512, 544
Central Family Planning Institute, New Delhi, 306, 347
Cernada, George P., 434-436, 459, 494, 501, 553
Ceylon, 59-60, 84-86, 398
Chai, John J., 266, 295
Chandrasekaran, C., 7-8, 14, 20, 29-33, 47, 52, 62, 89, 150, 160, 176, 211, 216, 243, 268, 376
Chang, C. I., 337, 347
Chang, C. T., 386, 390, 394-400, 403-404, 410-411, 415-418, 424
Chang, M. C., 179-180, 211, 234-237, 240-243, 451-453, 467, 475, 489, 498, 501-503, 540, 553
Charbonneau, H., 131
Chao, John, 551, 554
Chen, F. L., 453, 501
Chen, H. C., 295, 475, 501-503
Chen, K. C., 52
Chen, Kwa Hwa, 27, 53, 265, 554
Chen, Peter, S., 399, 420-422
Chen, S. Y., 4, 15
Cheng, Mark C. E., 407, 424
Chi, I. C., 111, 128
Child health, 1, 216
Child mortality, 207, 254-256, 268-269, 278, 313, 486
Chile, 36
Chipman, S. S., 243

Chow, L. P., 33, 52, 128, 177, 211-213, 243, 347, 398, 424, 434-436, 447, 451-453, 457-460, 465-467, 501-503, 559
Chowdhury, M. A., 335, 350
Chua, S. C., 392, 424
Chung Bom Mo, 257, 293
Civil registration systems, 95, 544
Clague, A. S., 207, 213, 368, 377
Closed interval, 318. *See also* Birth intervals.
Coale, Ansely J., 6, 14, 195, 245, 295, 373, 377
Cochran, W. G., 217-218, 222-232, 243
Cohabitation, 57
Coitus, 70, 101, 109, 168, 184, 226, 311
Coitus-interruptus, 18, 336, 368, 509
Collver, A., 245, 295
Columbia, 4, 36
Committee for International Coordination of National Research in Demography (CICRED), 55
Community surveys, 19, 30, 324. *See also* Surveys
Conception, 103-104, 107, 112, 311
Condom, 18, 35, 42, 83; in specific countries, 35-36, 42, 64-66, 80, 83-84, 252, 432, 436, 439, 443-447, 465-466, 492-493; use-effectiveness of, 27, 61, 458-459, 497
Continuation rates, *see* Termination rates; *see also* under specific contraceptive methods
Contraceptive methods, 2-3, 164, 263, 303, 320, 368; *see also* specific methods; continuity of use of, 20-24, 57-59, 73, 83-86, 133-160, 170, 176, 187-190, 203, 314, 408-409, 532; duration of use of, 8-12, 51, 134, 150-153, 164-167, 176, 187, 190, 210, 335, 366, 509-510, 515-516, 532-533, 538-539; knowledge about, 307-310, 319-323, 328, 331, 335-340; medical complications, 24, 189; side effects, 51, 457; substitution of non-program methods, 19-24, 59, 73, 150-158, 189-190, 177, 207-208, 240, 255-257, 264-265, 321, 369, 510; births averted through, 193-194, 203, 210, 534, 539, 551; termination of use of, 60-64, 68-69, 86, 107, 134-135, 139, 152-157, 164-165, 187-189, 197, 208, 364, 409, 515-519
Contraceptive practice, 20-22, 123, 337, 351, 361, 509; attitudes towards, 11-12, 18, 264, 302-303, 307-309, 319-320, 327-340, 472-473, 546; births averted through, 88, 134, 164-166, 195-206, 313, 324, 510, 522-539; in specific countries, 77-88, 85, 196-197, 208, 417, 459, 465-469, 481-482, 495-496, 499, 536; births postponed by, 174, 318; extent of, 316, 319; open interval extended through, 318; overlap with postpartum amenorrhea, 86, 181-184, 197, 208, 420, 447, 510, 532; residual fecundability and, 103, 109, 127; socio-economic variables affecting, 5, 22, 35-36, 170, 316, 338-339, 545-546; variance in, 339-340, 523
Contraceptive supplies, 8, 165, 366, 370, 509
Conventional contraceptive methods, 18, 83, 191, 321, 336; continuation of, 170, 366; in specific countries, 61-62, 70, 73, 170; irregular practice of, 151, 509, 523; use-effectiveness of, 61
Cook, S. W., 324, 348
Coombs, L. C., 489, 502, 553
Corsa, L., 244, 503
Costa Rica, 36
Council for International Economic Development, 428, 502
Couple-years of production (CYP) index, 27, 70-71, 250-253, 526-527
Coupon system, 21, 24-25, 434-435, 443, 492; *see also* Taiwan family planning program
Cox, G. M., 218, 243
Cross, H. E., 131
Cuzzort, Ray P., 295

David, A. S., 13, 321, 348
Davis, Kingsley, 5-6, 141, 247, 295, 311-312, 319, 347, 351, 377, 427, 502
Dandekar, K., 100, 128
Death rate, *see* Mortality
Del Tufo, M. V., 392, 424
Densen, P., 131
Demographic simulation models, 351, 354, 363
Deutsch, M., 324, 348
Developed countries: population decline in, 4-5, 186, 315, 354; prenatal mortality in, 118; sample fertility surveys in, 234, 302-304

Diaphragm, 35-37, 50
Divorce, 96, 139, 152, 156, 164, 190-192, 197, 233-234, 240-241; age-related, 368, 375; influence on reproduction of, 207, 311
Dorn, H. F., 218, 243
Duncan, Beverly, 295
Duncan, Otis Dudley, 246-248, 257, 261-263, 275-278, 283-285, 291-296, 303-347
Duration per abortion, 191
Duration per birth, 166-167, 171, 176-178, 184-187, 196, 203, 208-209

Easterlin, R. A., 314-315, 347
Eaton, J. W., 130
ECAFE, 18, 23, 42, 45, 399, 422
Ecuador, 41, 169-170
Ecological fallacy, 247
Education: fertility related to, 5-6, 247, 250-253, 259-262, 276-278, 313, 321-322, 390, 418, 453; in specific countries, 399-403, 408, 439-447, 456-457, 467-468, 488-489, 497, 500; of acceptors, 23-24, 30, 34-38, 43, 48, 180, 308; of children, 254-255, 268-272, 278, 486, 253, 265-267; of women, 322-323, 332, 382-383
Educational opportunities, 261, 267
Egypt, 36, 41
Eligibility for family planning, 27-28, 40, 57, 366-369, 394, 398, 436-439, 511-512
Employment: acceptance and, 48, 226, 26-62, 289, 307, 332, 399, 432-433; agricultural labor, 48, 250, 253, 268, 278, 313, 383, 486; fishing, 48, 278, 383, 486; industrial labor, 383, 395, 428; of women, 2, 289, 321-322, 382-383
Environment, 246-247, 253, 311; simulation of, 352
Ernakulam (Kerala), 38
Erhardt, C. L., 114, 131
Ethnic characteristics, 262, 322; acceptance and, 34-37, 43, 170, 198-203, 206, 332, 393, 399-400; differential fertility and, 322, 384-390, 394, 406, 410-413; *see also* Singapore, ethnic groups
European fertility project, 245
Evaluation, *see* Family planning programs

Factor analysis, 277, 288-291
Family planning: clinics, 3, 8, 18, 22-24, 47, 187, 216, 250-251, 261-263, 295-296, 321, 331, 393, 406, 419-420, 510, 515-516; information, 1, 8, 216, 319-320, 419, 509; personnel, 3-4, 8, 22, 48-50, 87, 216, 221-222, 258, 263, 265, 294, 321, 323, 329, 331, 352, 366, 370, 420, 431, 434, 489, 492-494, 509, 515, 540; services, 4, 12, 55, 87, 216, 248, 314, 320, 324, 329, 331, 419, 421-423, 430, 432, 434, 476, 480, 482, 494-495; strategies, 319-321, 353, 370-373
Family planning programs: administration, 3-7, 11, 22, 47, 55, 87-88, 307, 352, 370, 508, 552-553; costs, 6-7, 48-49, 232, 252, 258, 263, 282, 314, 317, 352, 360, 370, 420, 432-434, 492-493, 508; duration, 8, 86, 512; evaluation of, 7-13, 17-20, 24-25, 42-51, 98, 134, 187, 191, 216-242, 245, 265, 301, 306-341, 346-394, 459, 467, 501, 506-508, 546, 552-553; funding, 37, 81-88, 216, 307, 395; long-term planning, 11, 176, 253; mass media and, 50, 320, 395, 419-420, 488-489, 494; objectives, 1-8, 11-13, 17, 28, 49-50, 59, 195, 215-216, 221, 393, 399, 434-435, 493, 506; post partum, 33-34, 48, 93, 181-184, 395-397, 406-409, 414-415, 446-447; *see also* specific countries
Family size, 1-2, 57, 366; completed, 94; desired, 23, 28, 126, 226, 253, 264, 301, 351, 363-368, 402, 435, 531; effect of son preference on, 34, 402, 435; limitation of, 142, 152, 156-158, 189, 237, 318, 375, 401-402, 416, 423, 439-442, 447, 456, 473, 521; social norms regarding, 226, 237, 242, 264, 303, 307-309, 312-313, 318-320, 335, 339, 401.
Fawcett, James T., 257, 296, 314, 347
Featherman David, 295
Fecundability, 96, 142, 165-166; of acceptors, 10-11, 174-177, 358; decrease in, 374, 521; effect of age on, 108, 168, 174, 355, 374-376, 531; effect of contraception on, 109, 127, 133, 167, 196, 311, 315, 355-358, 365, 532; frequency of intercourse and, 70, 101, 109, 168, 184, 226, 311; increase in, 108; methods of estimating, 102-107, 180-181; natural level of, 103-107, 127,

311, 315, 365-366, 528; physiological factors in, 101-105, 118-119, 168, 363-365; simulation model for, 355-359, 363-366; variation among women, 365, 375, 529-532
Fecundity, 94, 98-99, 105, 124-126, 166, 174, 206-207, 239-240, 361
Feinberg, W. E., 378
Fertility behavior, 12-13, 321; theories of, 93-99, 177, 248, 257, 294, 528; variation in, 226, 339-340, 541, 544
Fertility decline, 3, 216, 317-318, 367, 383-388 409-418, 430, 469, 476-495, 501, 548-550; in industrialized countries, 4-5, 186, 315, 354
Fertility determinants, 131, 257, 262, 527-539; behavioral, 103, 108-109, 245, 259, 264-265, 302, 314, 322, 338-339, 351; biological, 66, 351, 355, 359-360, 363, 367, 371; demographic, 65-67, 93, 245-249, 276, 359, 366, 371, 386-390, 410-413, 417-418, 430, 478-486, 500; program, 4-13, 50-51, 65-67, 195, 217, 233, 249, 254-255, 315-339, 359, 412-420, 507-552; socio-economic, 2, 5-6, 13, 50-51, 61, 128, 245-247, 254, 259, 291, 311-315, 321-322, 338, 351, 390, 418, 435, 480, 485-489, 496-501, 507, 545-546, 550; theoretical formulation of, 257-261, 352
Fertility levels: areal variation in, 245, 248, 262, 276, 289, 292; natural, 93-102, 128, 527-529
Fertility rates, age-specific, 172, 263, 271, 278-281, 307, 286-288; age-specific marital, 69, 177-181, 194, 208, 278-280; decrease in, 68-69, 77-78, 102, 124-125, 164-168, 190, 327; estimation of, 10, 13, 69, 108, 171, 177-181, 194, 208, 231, 235-242, 245-248, 278-280, 289, 326-329, 542; in specific countries, 59, 61, 65-69, 73, 75-76, 79, 81, 179, 254-255, 386-390, 410-413, 417-418, 460, 465, 479-487, 495-499; potential, *see* Potential fertility
Fetal death, 98-102, 105, 113, 166, 226, 315, 355-356, 361, 376, 456; age of mother at, 116, 169, 185, 374-375; fecundability and, 102-103, 118-119, 363-364; incidence of, 112-118, 185, 359, 362, 374; effect on future reproduction, 207; pathological causes of, 118, 311; pregnancy rank and, 117; registration of, 112
Fiji, 23
First birth, 57, 58, 108-109
Follow-up studies, 12, 24, 74, 134-135, 145, 152-153, 159, 188, 234-235, 240, 292, 467, 510-511, 515-518; *see also* Taichung IUD follow-up survey
Forrest, Jacqueline, 52, 503
Frank, L. Z., 350
Freedman, Ronald, 5, 14, 28, 41, 48, 52, 134, 139, 153, 160, 170, 211-213, 221-222, 243, 257, 293, 296, 308, 316, 332, 347-348, 419, 424, 430-434, 457, 471-473, 485-490, 501-503, 511, 514, 524, 540, 547, 553
French, F. E., 112-115, 128, 131
Fresel-Lezey, M., 131
Freymann, Moye W., 78, 216, 243
Friedlander, Stanley, 245, 296
Fruchter, Benjamin, 289, 296

Ganiage, J., 131
Gautier, E., 131
Gendell, Murray, 213, 359, 378
Generational span, lengthening of, 65
George, A., 211
Georgia (USA), 240-241
Germain, Adrienne, 52
Gestation, *see* Pregnancy
Ghana, 4, 60, 84, 87
Giesbrecht, F. G., 377
Gillespie, R., 492, 502
Gini, Corrado, 103-106, 129
Girad, A., 5, 14
Goldberg, Arthur S., 259, 296
Goldberger, Arthur S., 291, 296
Gopalswami, R. A., 192, 211
Gordon, J. E., 99, 130-131, 213, 540, 554
Gordon, Robert A., 255, 291, 296
Grant, James, 429, 502
Great Britain, 302, 315, 550
Greenberg, B. G., 111, 128, 221-225, 237, 240, 243
Grizzle, J. E., 229, 243
Gynecological lesions, as cause of sterility, 126

Haller, Archibald O., 296
Hammes, L. M., 101, 129
Hannan, Michael T., 248, 296
Hardee, Gilbert, 333, 348

Harmon, Harry H., 289, 296
Harris, M. E., 243
Hashmi, S. S., 337, 348
Hauser, P. M., 5, 14, 335, 348, 427, 502
Hauser, Robert M., 291, 296
Hay, C., 347
Haynes, M. A., 192, 211
Health: indicators, 428; practices, 216, 430, 480; services, 1-3, 5, 81, 216, 262, 274, 393, 430, 433
Heer, David M., 245, 296
Heise, David, 283, 296
Henripin, J., 106, 129-130
Henry, Louis, 57, 89, 95-96, 102, 105-108, 122-125, 129-131, 171, 186, 209, 212, 359, 363, 373, 377, 527
Hermalin, Albert I., 250-251, 254-257, 261, 265-273, 278, 286-287, 291-292, 295, 296, 321, 348, 424, 459, 485-489, 500-502, 505, 544
Hertig, A. T., 118, 129
High fertility groups, 5, 419, 486
Hill, Stycos and Back, 223, 324, 348
Historical demographic data, 125-126, 245
Hodge, Robert, 266, 298
Holmberg, Ingvar, 254, 363, 371, 375, 377
Hong Kong, 42, 179
Hong Kong Family Planning Association, 20, 23, 30-31, 35-38, 40, 42, 52
Honolulu, 289
Horvitz, D. G., 111, 128, 366, 377
Hospitals, Family planning services of, 3, 18
Household composition, 322, 432
Household registers, 234-236, 263, 432, 512
Household size, 326, 433
Household surveys, 327-329, 335
Hsu, S. C., 398, 424
Huguet, J., 347
Husbands, surveys of, 324-325
Hyrenius, H., 131
Hysterectomy, 100

Illegitimate births, 96, 255, 280, 483
Immerwahr, G. E., 211
Impotence, 311
Incentives, 2-4, 47-48, 432, 492-494
Income, 23, 307, 318, 399-400, 403, 429; fertility related to, 250-252, 262, 313-314. 390
India, 2-3, 47, 61-62, 84, 186, 325-326
Indian family planning program, 249, 310, 316; acceptors, 23, 34, 38, 60-62, 84, 87; IUD, 61, 84, 249, 310; sterilization, 3, 36-38, 61-62, 84, 192-193, 249
Indianapolis Fertility Survey, 302
Individual client record, 21-25, 34, 509-510
Indonesia, 20, 23, 62, 84, 87
Infant mortality, 119, 121-122, 260, 356, 361, 368; decline in, 1, 88, 247, 383, 430; effect on fertility of, 119-120, 292-293; rate, 262
Infertility, 123-126
Intercourse, *See* Coitus
International Bank for Reconstruction and Development (BRD), 62, 89
International Planned Parenthood Federation, 406, 419
International Union for the Scientific Study of Population (IUSSP), 319, 322-324
Intra-uterine device (IUD), 18, 27, 35-39, 42, 51, 83, 509; acceptance rates for, 18-20, 24-26, 287; continuation rates, 61-69, 135-141, 144-150, 196-203, 448-459, 493, 497-498; expulsion of, 141, 151, 189, 197-202, 455-456 ; in specific countries, 24-25, 27-29, 32-33, 35-38, 42, 61-62, 64-65, 68, 70, 74, 78, 83-87, 195, 249, 252-255, 286, 310, 409, 414, 431-438, 448-465, 470-475, 490-498; Lippes loop, 20, 39, 59, 64, 431, 436-438, 456-459, 464, 492-493, 497; medical removal of, 141, 197-202, 455-456; ota ring, 156-158, 456, 464, 524; *in situ* during pregnancy, 135-137, 141-142, 153, 196-202, 450, 454-457, 515, 521; reinsertion of, 20, 150-153, 436, 450, 454, 463, 471, 523; side-effetcs of, 220, 395; sterility and, 197; substitution of, 192-193, 197, 317, 366, 454, 468, 523; termination rates, 59, 61, 64-66, 134-143, 153, 188, 366, 443, 452-457, 468, 472, 497
Intra-uterine mortality, *see* Fetal death
Iran, 36, 41
Isbister, J., 171, 195-196, 207, 212, 368, 535

Jain, Anrudh K., 88-89, 106-109, 129, 131, 177, 212, 447, 450, 453, 456, 470-471, 501-503, 514

Jain, Sagar C., 60, 89
Jahoda, M., 324, 348
Jamaica, 63-65, 84, 87
Japan, 5, 75, 192, 421
Jacquard, Albert, 352, 363, 371, 375-377
Java, 62
Johnson, C. T., 239, 244, 255, 260, 266, 285, 297
Jones, E., 131
Jones, S. B., 60, 89
Jöreskog, K. G., 290, 297, 298

KAP surveys, 242, 265, 293, 303-339, 342-345, 434-435, 452, 465, 544-567; *see also* Surveys
Karkal, Malini, 20, 52, 150, 160, 164, 211
Keeney, S. M., 3, 14
Kelley, W. J., 188, 207, 211
Kendall, Patricia L., 263, 297
Kenya, 36-38
Kerala (India), sterilization program in, 192
Keyfitz, Nathan, 212, 373, 377
Khanna, 540-541
Khoo Chain Kim, 408, 424
Kimpo, 310, 332
Kiser, C. V., 302, 348
Kitagawa, Evelyn, 263, 297, 412, 424
Knodel, J., 131
Korea, 65-67, 84, 87-88, 398
Korean family planning program, 20, 34, 64-67, 84, 87, 151-152, 195-196, 475; acceptors, 21, 30-31, 38; contraceptive methods in, 27-29, 64-66, 83-84, 87, 195, 252-254, 310; surveys, 30, 65, 310, 333
Koyang, 310, 318, 333
Krishnamoorthy, S., 350
Krotki, K., 348
Kuder, Katherine, 33, 47, 52
Kurup, R. S., 214
Kwa, S. B., 407, 408, 424

Lachenbruch, P. A., 377
Lactation, 306, 394, 447; duration of, 62, 88, 510; contraceptive function of, 88, 336; *see also* Breast-feeding
Laing, John E., 73, 89
Land, Kenneth C., 278, 297
Lapham, R., 7, 14, 48, 52, 508, 553
Latin America, 2, 179, 304-305
Lawley, D. N., 291, 297
Lazarsfeld, Paul, 263, 297
Le Bras, Hervé, 289, 297
Lee, B. M., 171, 195-196, 207, 212, 367, 535
Leridon, Henri, 128, 131, 353, 371, 377, 528-530, 533
Lewis-Faning, E., 302-303, 315, 348, 550
Lewit, Sarah, 20, 53, 134, 152, 161, 193, 214, 220, 244, 540, 554
Life expectancy, 351, 428
Life Table analysis: of births averted, 187; of fetal deaths, 112-116; of use-effectiveness of contraception, 133-145, 149-152, 159-164, 184, 187-190, 197-210, 450-451, 457-458, 473, 514-526, 552; models, 351, 521
Light, R. L., 216, 244
Lim Chong Ya, 383, 425
Lin, T. H., 243
Lingner, J. W., 130, 213, 363, 378
Linn, Robert L., 298
Liu, T. H., 144, 150, 160, 211, 451-453, 467, 501, 521-523, 553
Liu, P., 297, 503
Live births, 1, 101, 112-113, 260, 325, 337, 355-356, 364; non-susceptible period following, 101-102, 168, 193, 355, 358-359, 364-365, 374; registration of, 123, 236-237
Live birth rates, 164
Livi-Bacci, Massimo, 245, 297
Loewenthal, N. H., 312-313, 322, 348
Lorimer, Frank, 131, 245, 297
Low fertility areas, 419, 486
Lu, L., 494, 501

MacCorquodale, Donald W., 513, 553
MacKinnon, C. F., 243
McKusick, V. A., 131
Madura, 62
Malaya, *A Report on the 1947 Census of the Population*, 392; *see also* Western Malaysia
Malthus, Thomas, 103
Marcoux, Alain, 79, 90
Markov chains, 150, 153-160, 524-526
Marks, Eli S., 303, 348
Marriage, 374; age at, 59, 64-67, 74-75, 95-96, 125, 128, 278, 283-284, 361, 368, 486; dissolution of, 70, 96, 139, 152, 156, 164, 190-192, 197, 233-234, 240-241, 311, 368, 375; duration of, 315, 337, 362, 489; early, 95-96, 390, 392; first births in, 57-58, 109; fertility

within, 5, 28-29, 59, 62-67, 71, 83, 250, 254-255, 259, 280, 351, 373, 392, 410-413, 421; in non-contracepting societies, 93-98, 104, 128, 233; informal, 179; late, 65-67, 368, 382, 385, 390-392, 404, 410-413, 418, 421, 478-483, 495; percent married, 56, 67-68, 204, 254, 278, 318, 486; periods of separation in, 127, 311; status of, 56, 67-68, 234, 262, 307, 366, 433, 529; trends in, 64, 74
Married women of reproductive age, 29, 48, 56, 60, 94-96, 103, 108-169
Marshall, J., 128
Masnick, Stephan, 129, 209, 213, 359, 364, 378
Mass education, 322
Mass media, 50, 62, 260, 320, 395, 419-420, 488-489, 494
Mauldin, W. Parker, 5-7, 14-15, 48, 52, 143-147, 160, 212, 335, 348-350, 453, 459, 480, 503, 508, 521-523, 553
Matching studies, 10, 13, 179-180, 215-242, 541-542
Mayer, A. J., 131
Medical practice, 21-22
Menken, Jane, 98, 130, 213, 318, 349, 355, 359, 363, 367, 373, 378, 528, 537
Menopause: average age at, 99-100, 173; termination of contraception at, 64, 139, 156, 168, 375
Menstruation, 24, 88, 99, 168, 306, 318, 510
Method-acceptor, 19, 261, 508-509
Migration, 48, 262, 366, 382
Miscarriage, 18, 101-102, 104, 108, 117-119, 355, 364-365
Mishler, E. G., 214
Mohapatra, P. S., 197, 212
Mömmlingen, 121
Moots, B., 156, 160, 543-554, 503
Morgan, J., 332, 347, 349
Morocco, 67-69, 84-85
Mortality, 189, 208, 262, 276, 374; age at, 95, 366-368, 375; of acceptors, 61, 64, 71, 84, 156, 164-165, 170, 189-192, 197; of children, 207, 254-255, 268-269, 278, 313, 486; of infants, 1, 88, 119-120, 247, 260, 292-293, 356, 361, 368; registration of, 302, 432-433
Mortality rate, 5, 9, 65-67, 96, 226-227, 276-278, 311, 428; age-specific, 189, 197, 351, 366; decline in, 5-6, 65, 383, 422, 428-430, 486
Mosteller, F., 244
Mukerji, S., 181, 212
Multiple regression techniques, 241, 246, 251-256, 268-277, 285-288, 486, 542-544
Multiple classification analysis, 331-333
Multiple segment analysis of contraceptive use-effectiveness, 83, 150-151, 156, 523
Multivariate analysis, 10, 13, 246-249
Multivariate areal analysis, *see:* Areal analysis
Murty, D. V. R., 211, 276
Muthiah, A., 350
Mysore population study, 192, 316

Namboothri, D. N., 52
Nayer, P. S., 211
Nepal, 20, 73
Nerlove, Marc, 245, 258, 264, 276, 292, 297
New, M. L., 131, 213
Noe, L. F., 348
Non-acceptors, 88, 94, 231-233, 257
Non-contracepting populations: natural fecundability of, 93-98, 104; sterility and, 118, 124-126, 186; variance in fertility of, 128, 233
Non-program family planning, 3, 17, 21, 57-59, 66, 189, 237-239, 351, 546; in specific countries, 165-166, 170, 398, 406-407, 419, 422, 465-468, 480-483, 486, 498, 500-501; substitution of program methods for, 255-257, 264-265, 321, 510, 539
Non-susceptible period, 11, 93, 96, 105-107, 124, 169, 193, 355, 358-359, 374, 530-532, 539; *see also* Post-partum amenorrhea
North Carolina family planning survey, 337
Northwestern Europe, lowered fertility in, 5
Nortman, Dorothy, 3, 15, 36, 41-42, 52, 143, 212, 321, 348, 507-511, 553
Notestein, F. W., 6, 15
Nuclear families, 324
Nuptiality, *see* Marriage
N'zakara (Central African Republic), 118

Observational studies, 217, 222, 224, 232-234, 241
Okada, L. M., 234-236, 240, 243
Open interval: acceptance and, 34, 43, 190, 239-241, 333, 439, 446-447, 467-468, 497; fertility change indicated through, 318, 367, 374, 550; lengthening with contraceptive use, 318
Oral contraceptives (pill), 18, 27, 34-38, 42, 68, 83, 170, 337, 395; in specific countries, 25, 29, 34-37, 42, 62, 65-66, 68, 73, 78, 80, 83-84, 252, 395, 408-409, 431, 434-436, 439, 442-443, 447, 457-458, 465-466, 497; irregular practice of, 151, 509, 523; use-effectiveness of, 447, 457-458, 475, 497
Ovulation, 99-101, 103, 107, 446-447
Ow Chwee Huay, 383, 425

Pakistan, 70, 85, 90, 333-336; family planning program, 34, 70-71, 85, 314
Parity, 95, 125, 226, 231, 361, 366; related to acceptance, 22-23, 29-43, 57, 83, 142-143, 167, 172, 177-180, 207, 361-368, 399-402, 439-443, 455-457, 467, 489, 497; related to age, 257
Park Chai Bin, 289, 297
Parker, M. P., 107, 130
Parthasarathy, N. R., 214
Parturition, 254, 311
Pascal, J., 120, 129
Patankar, T., 177, 213
Path analysis, 278-285, 542
Peabody, Gerald R., 377
Pearl's index, 127, 137, 514
Peng, J. Y., 503
Perera, T., 59, 90
Perez, A., 120. 129
Perrin, E. B., 207, 315, 349, 363-364
Philippines, 71-74; family planning program, 20, 23, 31, 73-75, 83, 85
Planned Parenthood Federation of America, 52
Poole, Kenneth, 376-377
Population: census, 29-30, 65, 71, 179, 263-266, 289, 301, 338, 381, 385, 392, 512, 544; density, 262, 381, 422; growth, 2-3, 70, 81, 354, 370, 381-385, 390-393, 402, 413, 421; heterogeneity of, 141, 148, 181, 328, 369, 371; parameters, 326, 353; policies, 2-3, 49, 84, 315, 376, 395; (in specific countries, 2, 6, 62, 79, 81-82, 84, 393-397, 408, 421-423, 430); records, 237; reduction, 78, 84-86, 216, 412-413, 419, 430, 476-477, 483-485, 496, 499, 550-551; (in specific countries, 57-62, 64-67, 70-79, 82, 84, 87, 317, 398, 398, 421, 431, 496); registers, 301; registration statistics, 197, 544; simulation model of, 352-354, 370; trends, 189, 359
Population Council, 322-323, 349, 433, 435, 480, 518, 540, 553
Population Growth Estimation Projects, 303
Portes, Alexandro, 296
Portugal, 245
Post-abortal amenorrhea, 190-191
Post-partum amenorrhea, 24, 65, 70, 82, 101-102, 120, 127, 203, 210, 226, 306, 315; acceptance during, 83, 181-184, 196-197, 204-205, 239-241, 364; duration of, 119-120, 123, 184, 197, 204-205, 209, 365, 374-376; overlap with contraceptive use, 86, 181-184, 197, 208, 420, 447, 510, 532
Post-partum anovulation, 142, 164-168, 172, 364-365
Potential fertility of acceptors, *see* Acceptors
Potter, Robert G., 86, 90, 107, 121-123, 129-130, 134, 148, 160, 171-176, 183-187, 191-193, 196-197, 203-209, 212-214, 321, 349, 359, 363-365, 368-370, 374, 377-378, 447, 450, 457, 475, 502-503, 514-523, 527-528
Pregnancy: acceptance during, 61, 156, 172; accidental, 316, 319, 420, 533, 539; attitudes and values related to, 303, 399; duration of, 101, 104, 113, 166, 190, 358, 363-365, 374-375; during contraceptive use, 133-137, 141-142, 153, 164, 184, 196-202, 450, 454-457, 515, 521; full term of, 101, 118, 123; risk of sterility in, 124, 209; short-lived, 118; *see also* Fetal death; termination of, 23-24, 325, 337, 359, 375; unwanted, 264, 406, 420, 435
Pregnancy history, 22, 306, 317-319, 325, 337
Pregnancy rank, 97, 116, 117
Pregnancy rates, 198-202, 206, 208-210, 235
Premarital conception, 104, 109

Prenatal health care, 118
Press, S. James, 258, 297
Private birth control, *see* Non-program family planning
Private family planning associations, 3, 17, 21, 406
Program-acceptor, 19-21, 26, 508-509
Prospective acceptor, 44-46
Provisional acceptor, 18, 509
Puberty, age at, 99-100
Public health practice, 21
Puerperium, 124
Puerto Rico, 188, 292, 324
Punjab, 122
Pullum, Thomas W., 513, 553

Quah Siam Tee, 407, 409, 424-425
Quebec, infant longevity in, 121
Questionnaires, 322-323, 336; content of, 306, 315-319, 325-326, 346; IUSSP model, 319, 322-323, 346; Population Council model, 322-323, 346; use of, 323, 339

Radick, Annette, 373, 378
Ram, Bali, 213
Randomized responses techniques, *see* Surveys
Random sampling procedures, 326
Ravenholt, R. T., 551, 554
Regional surveys, 303
Rele, J. R., 177, 213
Religion, fertility and, 262, 305, 313, 322
Reproductive behavior, 351, 361; effect of contraception on, 353-354, 359, 362-365, 370; knowledge of, 303; variables affecting, 99-102, 117, 311-315, 351-352, 367, 375
Residence, fertility and, 250-254, 261-262, 314, 321-322, 390
Retel-Laurentin, A., 118, 130
Retrospective surveys, 12-13, 323-324, 515
Reynolds, J., 7, 15, 507, 554
Rhodes, Lewis, 245, 297
Rhythm, 18, 336-338, 509; in specific countries, 64, 73, 83
Rider, R. V., 128, 347
Ridker, R., 3, 15
Ridley, J. C., 130, 177, 207, 213, 353, 363, 367-368, 374-377
Riew, John, 250-253, 256, 297
Robinson, A. H., 247, 275
Robinson, W. S., 297
Rochat, R. W., 231, 240, 243
Rogers, E. M., 3, 15
Ross, John, 23, 29-34, 37, 41, 52, 90, 134, 152, 160, 177, 213, 297, 349, 475, 503, 528
Rossi, P. H., 7, 15
Royal Commission on Population, 302
Royaume du Maroc, 67, 90
Rubin, D. B., 223-224, 243
Rummel, R. J., 289, 297
Rural areas: acceptance in, 34, 43, 429, 431, 446, 456-457; surveys in, 303, 328-329, 336
Ryder, N. B., 417, 425

Sagi, P. C., 214
Sakoda, J. M., 365, 378
Sampling, *see* Surveys
Sarma, R. S. S., 366-367, 378
Saw Swee Hock, 381, 386-387, 421, 425
Schoenberg, Ronald, 266, 287, 291, 298
Schultz, T. Paul, 250-251, 253-257, 259, 261, 264-276, 276, 286, 292-294, 297-298, 312-314, 349, 500, 503
Schultz, Theodore W., 245, 298
Scrimshaw, S. C., 169, 213-214
Seasonability of births, 240
Seltzer, William, 7, 13-15, 70, 89, 194, 213, 348-349
Senegal, 121
Sex ratios, 262-263, 382-385, 412
Sexual taboos, 128
Shah, B. V., 377
Shapiro, S., 114, 131
Sharman, A., 184, 214
Sheps, Mindel C., 96, 130, 207, 213, 315, 318, 349, 351, 352-355, 359, 363-364, 367, 371-375, 377-378
Shoenbucher, A. K., 244
Siegel, Paul, 266, 298
Silver, Morris, 245, 296
Simmons, George, 249-252, 257, 261, 298
Simulation models: ACCOFERT, 374; BIOSIM, 355-359, 361-363; computer models, 180-181, 206, 210, 360-361, 367, 372-376; FERMOD, 365, 374; POPREP, 375; POPSIM, 366, 373-375, 538; REPSIM, 367, 373-375; of acceptance, 177, 358, 362; of birth rates, 352, 358-359, 366; of births averted

164, 174, 207, 210, 368-370, 374; and family planning programs, 353-376; of fecundability, 355-358, 363-366; of fertility trends, 352, 359, 364, 372; of open interval, 367; of population, 352-354, 370; of pregnancy, 355-358; demographic variables of, 351, 354, 359-360, 363, 366, 369, 371-373; limitations of, 354, 361, 366, 371-372, 374

Singapore, 381-424; differential fertility in, 317-318, 384-390, 410-413, 419; ethnic groups in, 34, 37, 198-201, 203, 206, 382-393, 399-406, 410-413; fertility change in, 383-385, 388, 394-395, 398, 410-413, 415-416, 419, 421; illegal abortion in, 407; non-program contraception in, 170, 398, 406-407, 419, 422n; population policies of, 393-394, 396-397, 421-423.

Singapore Family Planning Association (SFPA), 35, 41, 393-395, 404, 406

Singapore Family Planning and Population Board, 31, 35, 394, 406, 419

Singapore family planning program, 393-396, 398, 406-407, 412, 418-424; abortion under, 407-408; acceptors, 23, 29-31, 34-37, 40-41, 191, 198-203, 206, 393-406, 408, 414, 419-423; births averted by, 206, 417, 536; births postponed by, 416-417; condoms, 35-36, 42; continuation rates, 198-201, 203, 409; conventional contraceptives, 170; incentives, 397; IUD, 33, 35-37, 42, 395-397, 409, 414; oral contraceptives (pill), 29, 34-35, 395, 408-409; postpartum programs, 33-34, 395-396, 406, 409, 414-415; sterilization, 407-408; surveys, 392, 397, 408-409, 414-415, 422

Singh, R. N., 148

Singur, 540-541

Sirageldin, Ismail, 312-314, 317, 348-349

Smith, D., 65, 91

Smith, T. E., 131

Social action programs, 4, 216-220, 222

Socio-economic development, 4, 87-88, 226, 247, 260, 291, 308; effect on fertility of, 2, 5-6, 13, 50-51, 61, 128, 245-247, 254-255, 259, 291, 311-315, 321-322, 338, 351, 507, 545-546, 550; in specific countries, 382-383, 390, 404, 418-419, 428-429, 435, 446, 480, 485-489, 495-501; threshold hypothesis of, 5-6

Socio-economic strata, 338; differential contraceptive behavior of, 170, 190, 308, 316, 320

Socio-economic variables, 262-264, 307, 321-322

Som, R. K., 337, 349

Sonquist, J., 332, 347-349

Spacing of births, 11, 23, 142, 152, 189, 318, 364-365, 375, 399-402, 416-423, 435, 439-444, 447, 456, 473-474, 550

Speare, A., 298

Sperm count, effect on fertility of, 101, 168

Sri Lanka, 20, 23

Srikantan, K. S., 308, 349, 552

Srinivasan, K., 211, 318, 325, 349-350, 376

Stanley, Julian C., 541-542, 553

Starmer, C. F., 243

Stephan, Frederick F., 52, 143, 160, 212, 553

Sterility, 11, 28, 105, 207, 361, 394; age-related, 102, 125-126, 169, 173, 177, 186, 368, 374-375; births averted through, 174, 185; in men, 186; incidence of, 315-316; monthly probability of, 166, 185, 204, 209, 315; in non-contracepting populations, 124-126, 186; pathological reasons for, 118, 168-169, 174, 186; pregnancy-related, 124, 309; primary, 124-125, 173; program acceptors and, 65, 70-71, 164, 172, 174, 176-177, 185, 196, 204-205, 208; resulting from abortion and miscarriage, 117, 124, 174; resulting from contraceptive use, 316; secondary, 124-125, 197, 531-532, 539; socio-cultural determinants of, 173; temporary, 168

Sterility rates, 185-186, 315

Sterilization, 3-4, 18, 35-36, 42, 51, 167, 320, 361, 368, 454, 509; in specific countries, 25, 36, 61-62, 66-67, 70, 80, 84, 139, 152-153, 156-159, 192-195, 204, 249, 308, 366, 407-408, 454, 471-472, 487

Still-birth, 112, 118, 190, 365, 374; registration of, 123

Stoekel, J., 335, 350

Su, K. H., 473

Sub-fecundity, 126
Substitutions, *see* Contraceptive methods; Taiwan family planning program
Suchman, E. A., 4, 7, 15, 506, 554
Sullivan, J. M., 503
Sun, T. H., 88-89, 424, 430, 447, 450, 453, 456, 470-471, 501-503, 553
Sundong Gu, 310
Surveys, 306-329, 338-345; administration of, 328-329, 333-337, 339-340; analysis of, 307-312, 323, 329-335, 339-340; data collected, 12-17, 21-22, 110, 117-118, 169, 179-181, 185, 194, 302-303, 307-308, 310-312, 317-318, 320-324, 327-332, 337-340, 353; demographic, 303-305, 339; design of, 311-317, 321, 323-326, 330, 340; errors and biases in, 318, 335-337, 341, 510, 547; evaluative functions of, 98, 187, 191, 231, 245, 265, 301, 306-341, 546; in specific countries, 28, 61, 117, 122, 134, 138-144, 153-156, 192, 220-222, 302, 304-306, 308-310, 315-316, 318, 322, 331-333, 337-338, 355-356, 392, 397, 408-409, 414-415, 422, 428, 431-435, 452-457, 460, 465, 485-487, 490-493, 496, 520-522, 540-541; KAP/fertility, 179, 242, 293, 303-345, 434-435, 452, 465-467, 454; methods, 304, 310, 324-326, 342-345; multipurpose, 303-305, 339; objectives of, 306-315, 317, 324, 326, 329, 339; pilot studies for, 327; questionnaires, 306, 315-319, 322-323, 325-326, 339, 346; randomized response techniques, 337-338; sampling 301, 303-304, 306-307, 324-329, 340; sponsorship of, 305-306, 338, 342-345

Taek Il Kim, 64, 90
Taichung family planning program, 177, 250, 431
Taichung IUD follow-up survey, 28, 134, 138-144, 153-156, 220-222, 308-309, 331-332, 452-457, 460, 490-492, 520-522, 540
Taiwan: aboriginal areas of, 255, 276, 486; abortion in, 77, 308, 455, 460, 470-472, 475, 487, 495-496; age-specific fertility in, 75-76, 254-255, 460, 465-467, 479-487, 495-496, 499; fertility decline in, 427, 429-430, 435, 475-501; household registration system, 237, 427, 432-433, 445, 467, 496; marriage in, 74-78, 431, 478-486, 489, 495; modernization in, 428-429, 446, 480, 486-487, 495-497, 500; mortality decline in, 428-430, 486; National Health Administration, 435; non-program contraception in, 76-77, 165-166, 170, 465-468, 480-483, 486, 498, 500-501; population policies of, 74-78, 84, 398, 431-432, 494-498
Taiwan family planning program, 47; acceptors, 28-33, 38, 40, 74-77, 85, 401, 431, 434-447, 451, 455-457, 460-468, 473-475, 480-483, 486-498, 500; acceptor targets, 26, 76-78, 85, 431, 436, 496; births averted estimates from, 77-78, 85, 196, 459, 465-469, 481-482, 495-496, 499; clinics, 434, 463; condom use; 83, 432, 436, 439, 443-445, 447, 458-459, 465-466, 492-493, 497; continuation rates, 448-457, 459, 493, 497-498; costs, 78, 432, 434, 456, 492-493; coupon system, 24-25, 31, 434-435, 443, 492; evaluation of, 434-435, 494; fieldworkers, 4, 26, 34, 45, 47-48, 74-78, 255, 265, 275-276, 294, 427, 429, 431-432, 434, 446, 456-457, 467, 476, 485-489, 492-494; IUD use, 24-25, 29, 32-33, 83, 139-142, 254-255, 310, 431, 434-447, 450-457, 459-465, 467-468, 470-475, 480, 486-487, 490-494, 497-498; objectives, 434-435, 493; oral contraceptives (pill), 25, 83, 431, 434-436, 439, 442-443, 447, 457-458, 465-466, 475, 497; private doctors and, 431-432, 489, 492; service statistics of, 21, 24, 180, 427, 434-435, 496-497; sterilization under, 25, 308, 454, 471-472, 487; substitution effects, 454, 468, 486, 498; termination rates, 443, 450-457, 468, 472-473, 497
Taiwan family planning survey, 310, 318, 337-338, 428, 434-435, 465, 485-489, 492-493, 496; *see also* Taichung IUD follow-up survey
Taiwan Population Studies Center, 26, 45, 53, 433-434, 490
Taiwan Provincial Committee on Family Planning, 53, 69, 74, 90, 298, 431, 435, 441-442, 462, 464, 504
Taiwan Provincial Department of Cultural Affairs, 504

Taiwan Provincial Institute of Family Planning, 298
Taiwan Provincial Government, Department of Civil Affairs, 298, 445
Taiwan Provincial Maternal and Child Health Institute, 433-434, 490
Takeshita, J. Y., 28, 48, 52, 134, 139, 153, 160, 170, 211, 231, 234-236, 240, 244, 293, 296, 308, 316, 332, 348, 424, 430-431, 434, 457, 471, 490, 502, 511, 514, 524, 540, 553
Talwar, P. P., 178, 214
Taylor, W. F., 114, 131
Termination rates, 59, 60-66, 68-69, 86, 107, 134-143, 152-157, 187-189, 197, 208, 364, 409, 443, 452-457, 468, 472, 497, 515-519; *see also* Contraceptive methods; Intrauterine device
Thailand, 78-79, 85
Thailand family planning program, 20, 23, 78-79, 85, 87, 90, 310
Theil, H., 248, 298
Thomas, Edwin M., 275, 298
Thonk Kah Leong, 395-396, 424
Tietze, Christopher, 20, 53, 120, 127, 130, 134, 152, 160-161, 193, 214, 220, 244, 409, 514-519, 540, 554
Tolba, Salah El-din, 374, 379
Total Fertility Rate (TFR), 81-82
Tourouvre, 121
Traditional cultures, 5
Treolar, A. E., 101, 129
Tubal ligation, 66
Tunisia, 79-80, 87
Tunisia family planning program, 38, 49, 68, 78-80, 83, 85
Turkey, 81-82, 85, 91
Turkish family planning program, 41, 49, 81-82, 85, 87
Tyler, C. W., 244

Underdeveloped countries, 6
United Nations, 2, 6-7, 15-18, 22, 25, 42, 53, 56, 91, 133, 161, 226, 244, 266, 302, 316, 319, 335, 350, 425
United States, 302, 306, 373, 379
United States Bureau of the Census, 303
United States National Center for Health Statistics, 373, 379
Unhannand *et al.*, 79, 91
University of Cairo, 374
University of Michigan: Population Studies Center, 272-274, 433-435, 490; Survey Research Center, 331
Unwanted births, 4, 253, 261, 266, 312, 406
Upper Volta, 118
Urban areas: acceptance in, 34, 43, 429, 446, 467, 476; surveys in, 303, 328-329, 336
Urban-rural fertility differentials, 250-254, 261-262, 390
Urbanization, 254, 261, 431, 446, 456-457, 488-489, 500

Valmary, P., 131
Van Ginneken, Jeroen, 52
Vasectomy, 3, 36-38, 64-66, 70-71, 252
Vela, P., 129
Venereal diseases, 126
Venkatacharya, K., 173-176, 181, 189, 192, 207-208, 212-214, 318, 350, 368, 374-376, 379
Vig, O. P., 349, 298
Villages, 328-329
Vincent, P., 106-108, 130, 181, 186, 209, 214
Vital statistics, 263, 266
Voluntary family planning organizations, 47

Wan Fook Kee, 408-409, 421, 425
Wang, C. M., 4, 15
Warner, S. L., 111, 130, 337, 350
Watson, Walter B., 52, 65, 88, 91, 160, 310, 348, 350
Weaning, age of, 120
Wells, H. Bradley, 540
Werts, Charles E., 291, 298
Western Europe, fertility decline in, 354
West Malaysia, 170, 214, 237, 309, 329, 350
West Malaysia family planning program, 18, 23, 36, 41, 170, 240, 509
West Malaysia Family Survey, 170, 309, 322, 329, 350
West Pakistan Family Planning Survey, 310
Westoff, C. E., 181, 214
Whelpton, P. K., 302, 348
Widowhood, 96, 173, 207, 375, 391-392; of acceptors, 139, 156, 164, 189, 192, 197
Williams, W., 7, 15
Winokur, H. S., 244
Wishik, S. M., 27, 53, 526, 554

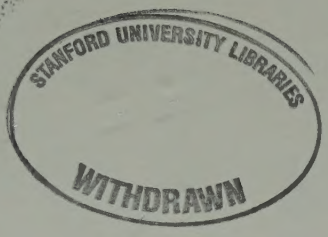
Wolfers, David, 33-34, 53, 96-98, 130, 169-174, 178-180, 185, 191-194, 205-206, 214, 368, 397, 408-409, 414-415, 422, 425, 528, 539, 550, 554
Wolfers, H., 170, 214, 397, 425
World Fertility Survey, 303, 307
World Health Organization, 130
World Population Conference, Bucharest, 5
World War II, fertility following, 383-385, 430
Wright, N. W., 59, 91
Wyon, J. B., 99, 130-131, 213, 540, 554

Yeh, Stephen, 397, 400, 403, 424
You Poh Seng, 381, 383, 392, 425

Zatuchni, G. I., 214
Zellner, Arnold, 298
Zucker, H., 131

Copyright © Organization for Economic Co-operation and Development (OECD)

D/1975/0019/9

Stanford University Libraries B47

36105036386683